DEPARTMENT OF MANAGEMENT SCIENCE

University of Strathclyde Business School

The carbon footprint of online digital performances

MSC. DATA ANALYTICS PROJECT

BY:

DIVJOT KAUR NARULA

STUDENT NO. 202083238

SUBMITTED TO:

DR. AINSLEY MILLER (ACADEMIC SUPERVISOR)

DR. FIONA MACLENNAN (INDUSTRIAL SUPERVISOR)

Table of Contents

Declaration

Except where explicitly stated, all the work in this dissertation – including any appendices – is my own and was carried out by me during my MSc course. It has not been submitted for assessment in any other context.

Signed -
Divjot Kaur Narula

Acknowledgements

I would like to take this opportunity to thank everyone who helped or supported me throughout my MSc. course and my summer project. I would like to thank Creative Carbon Scotland for giving me the opportunity to work on this interesting topic and providing me with relevant information and access to their data. Dr. Fiona MacLennan was very welcoming and prompt in responding to all my concerns and questions. I would also want to thank Civic Digits and Rupert Goodwins for explained the working of their shows and the internal technical details of production of the show, he answered all my silly doubts with extreme details. My project supervisor Dr. Ainsley Miller, who was a constant support since the start of my project and provided me direction and supervision all along my work. Everyone was very approachable, and this project would not have been possible without their help and guidance. Last but not the least, I want to thank my Mom who was a big moral support throughout my entire MSc. studies, unfortunately I lost her to COVID-19 before I could even finish my degree. My degree will be dedicated solely to her.

Chapter 1: Project Setting

1.1 Background of the Organization

Individuals and organizations interested in how environmental sustainability interacts with their work are encouraged to contact Creative Carbon Scotland. The firm was established in 2011 to root environment sustainability in the art and cultural sector of Scotland and ever since has developed enormously ("About Creative Carbon Scotland | Creative Carbon Scotland," 2021) . The Paris Agreement, which was recently adopted by the United Nations, demonstrates widespread recognition of this and what it means, both in terms of how culture must change in order to reduce climate crisis as much as possible and how it will inevitably change in response to the environmental changes that are already occurring ("About Creative Carbon Scotland | Creative Carbon Scotland," 2021). As a result, environmental sustainability is an issue that we believe Scotland's cultural sector cannot ignore and needs to be dealt with urgently. The organisation now supports nearly 120 theatre, art centers and touring organizations and also individual artists, craftspeople, producers, publishers, etc. to mandatorily measure, report and consequently reduce the carbon footprints of their artistic performances. As a repercussion of the pandemic, all live art performances have been suspended, which has led to many art organizations into adapting the online digital presentation of their productions. Thus, the focus is now on the role of this sector in transforming our society to address climate change. Information and communication technology (ICT) is mostly regarded as a greener alternative which helps substitute the physical processes into virtual ones in order to reduce the emissions to a large extent (Ong et al., 2014).

Civic Digits C.I.C is one such theatre company which has made groundbreaking progress in digital performances. The mission of the company is 'To create a digital future where we can all flourish'("About us – Civic Digits," 2021). They create creative learning opportunities for young people at the intersection of data, gaming, and performance, which are frequently presented in educational settings such as schools and universities, as well as theatres – where, for example, audiences are expected to interact with the Civic Digits website to download a phone app created for the show. 'The Big Data Show' created by them with the creative minds of award-winning playwright and theatre director Clare Duffy and a renowned tech-journalist Rupert Goodwins, is a digital and interactive data show designed to educate young people in the United Kingdom about cyber vulnerability ("The Big Data Show – Civic Digits," 2021). It addresses the growing threats to the well-being of young people as they navigate the digital world. In June 2020, they were supposed to conduct the first play in live venues which was already involving the use of IT for smartphone apps, cloud storage and the use of big data. However, with the outbreak of the COVID-19 pandemic they had to now produce it as a fully digital experience. Furthermore, as the theatres gradually open, their plan is to create online digital work to supplement live presentations. As discussed with the Civic Digits staff members, The Big Data Show will be presented in a number of mid-sized theatres to audiences ranging from 450 to 600 people. The performance will last 90 minutes, and the stage set will be marginal once again (dominated by a large screen), and the show will feature a presenting team of 5-6 actors, a presenter and technicians.

Future plans

According to the information provided in the client meeting with Creative Carbon Scotland and Civic Digits, during the pre-pandemic times, Civic Digits toured a version of the show around schools to 12 venues in the Perth/Edinburgh/Glasgow area. They used a minimal stage set with 5 people travelling in 2 vans and a car. In the future, they intend to give 83 performances to 2400 students across Scotland, from the Shetlands to the Borders. It will employ a similar set-up of people and transportation. They hope to present a parallel workshop at each location using a local presenter. Thus, using some technical equipment that adds to the carbon emissions by a margin. The company sometimes supplies up to 25 iPhones to schools which have to be shipped to the location ahead of the show.

1.2 Aim of the Project

The purpose of this paper is to evaluate face-to-face performances as a functional unit in order to define the relevant factors of power consumption and environmental impact of live performance production and compare it to other forms of digital performance such as zoom lectures, digital shows, video conferences and meetings, and so on. This project is a comprehensive study to evaluate the actual energy consumption which contributes the highest towards the carbon emissions in a streamed video performance as compare to a live theatre performance. The concept of carbon footprint extends beyond the use of electricity by products. It includes greenhouse gas emissions from energy and materials used throughout a product's life cycle.

The operating and lifespan energy consumption of end users, video streaming equipment, and network infrastructure (both fixed and mobile) are all part of the research. The energy consumed by audiences to attend these events is then compared, including direct fuel consumption of vehicles, public transportation (local and international), theatre or art center infrastructure, and number of travellers. We also investigate how differences in travel distance and video streaming duration affect the overall carbon savings generated by digital performances. Going forward with some initial cleaning and exploratory analysis with the given data of emissions in live face-to-face performances in Scotland in the year 2018-19 the main focus of the project is to scope the most important factors contributing to the carbon footprint of a digital streamed performance or a live performance. Electricity consumption can be seen as an important factor which in turn depends on the size of the organisation i.e., the number of employees and members of staff in that company, the given office floor area which is dependent on the type of the organisation.

Thus, keeping these basic things in mind and further doing some literature research we will scope out the irrelevant or non-impacting features from the data and build a regression model with necessary assumptions which can help us calculate the factors on which the most important factor of emissions depends. It has been evident that a large portion of the carbon emissions is linked back to the electricity consumption, so a majority of the ICT industry and its users are switching to renewable energy sources. The main assessment and analysis would circle around the fact that the energy consumptions in a digital performance and subsequent carbon emissions are infinitesimally small and thus, in the future lowering the carbon footprints. Previous investigations have shown that, in the worst-case ICT industry will consume as much as 51% of the earth's electricity in 2030 if there is not enough energy efficient wireless networks and data centers established. Also, ICT's electricity consumption would be responsible for up to 23% of global greenhouse gas (GHG) emissions (Andrae and Edler, 2015).

1.3 Proposed Plan

ICT impacts the climates in three popular ways. First is the direct carbon footprint from the emissions during production use and disposal. Second is the indirect positive or negative effects from using it, example- travel substitution. Third is impacting behaviours and preferences. Analysing the sensitivity of energy and carbon costs to various factors, as well as trends in energy and carbon consumption, to forecast how the comparison might change in the future (Ericsson, 2020).

Civic Digits produced 'The Big Data Show' and it is intended to be presented in both formats in the future so will be a key element in assessing and comparing the impact of live and digital presentations. With the help of some reports that have already done the analysis on ICT processes in general and understanding the technical details of 'The Big Data Show', an attempt is made to estimate the ICT footprint of digital event and then used the model obtained from live events to find inferences on lowering the annual emissions.

1.4 Project Schedule

Week 1	**Week 2**	**Week3**	**Week 4**	**Week 5**	**Week 6**
14-Jun	21-Jun	28-Jun	05-Jul	12-Jul	19-Jul
-Initial Research on CCS Website -Ericsson Paper -Getting touch with Fiona for data	-Another meeting with CIVIC Digits and understanding the data - Teams Data sharing and connecting with Rupert, Robyn and Claire	First Offical Meeting with Ainsley regarding using regression and Google Scholar	Getting in touch with Rupert on important factors for digital performance carbon emmission	Complete initial Data Analysis and Project setting part of the Report Meeting with CIVIC Digits and Fiona to understand production data for Online performances	Start with the digital performance key factors and work to find a relation
-Introduction meeting with Fiona and Ainsley	Doing some Literature Review on Project setting	Some Data Related Discussions with Fiona	Checking with Fiona on the size of organisation and employees or artist data	Complete Project Setting and Context Complete the data modelling for live performance and literature review for digital performance	Begin with data simulation Continue Literature review to find digital performance key factors Start with Structuring the Report
Week 7	**Week 8**	**Week 9**	**Week 10**	**Week 11**	**Week 12**
26-Jul	02-Aug	09-Aug	16-Aug	23-Aug	30-Aug
Complete Simulation and begin testing for digital performances Find out annual data of organisations	Finish all coding and modelling part and begin to finalise the client report Third Official Meeting with Ainsley	Writing the Client Report with further research	Send the 1st part of draft to Ainsley for review	Final Draft submission for Review	Final Hand-in Submission

Figure 1: Project Plan and Schedule

1.5 References

About Creative Carbon Scotland | Creative Carbon Scotland [WWW Document], 2021. URL https://www.creativecarbonscotland.com/about/ (accessed 8.9.21).

About us – Civic Digits [WWW Document], 2021. URL https://civicdigits.com/about-us/ (accessed 8.9.21).

Andrae, A., Edler, T., 2015. On Global Electricity Usage of Communication Technology: Trends to 2030, Challenges 6, 117–157.

Ericsson 2020. A quick guide to your digital carbon footprint carbon footprint, Ericsson, 2020.

Ong, D., Moors, T., Sivaraman, V., 2014. Computer Communications 50, Comparison of the energy, carbon and time costs of videoconferencing and in-person meetings, 86–94.

The Big Data Show – Civic Digits [WWW Document], 2021. URL https://civicdigits.com/project/the-big-data-show/ (accessed 8.9.

Chapter 2: Client Report

Executive Summary

People have been confined to their homes for more than a year due to the ongoing pandemic, which has prevented them from doing everything including studies to corporate jobs, and from conducting interviews to watching movies. All businesses will eventually convert their daily operations to a digital format causing increased use of Information and communication technology (ICT), which is viewed as an eco-friendlier approach towards the most critical and concerning topic of global climate change (Ericsson, 2020). Creative Carbon Scotland is an organisation that assists over 120 businesses in measuring and reporting their carbon footprints. Civic Digits, a digital theatre company, is one of them. Civic Digits have recently launched “The Big Data Show”, an educating show for the young generation to make them aware about cyber vulnerability in the form of an interactive gaming show (“The Big Data Show – Civic Digits,” 2021). As a part of this project, we are going to analyzed carbon footprint data, reported to the Creative Carbon Scotland by more than 100 art performing organisations in the year 2018-19, for the in-person live theatre performances and comparing to the new digital way of presenting these art forms. We can then compare the difference in carbon emissions between the two forms using literature and some technical details from Civic Digits about the functions and operations of an online performance.

Carbon footprint of an ICT impacts the climates in three popular ways. First is the direct carbon footprint from the emissions during production use and disposal. The second factor to consider is the indirect positive or negative effects of using it, such as travel substitution. The third aspect is influencing behaviors and preferences (Ericsson, 2020). Analyzing the sensitivity of energy and carbon costs to various factors, as well as trends in energy and carbon consumption, in order to forecast how the comparison may change in the future. (Ericsson, 2020). For the ease of our study and the limitation of being able to find details in a data that is reported just by a survey conducted in the organisations we have only considered the emissions from direct production use and disposal and the travel substitution components in the comparison to an in-person performance.

The first part of the project was to identify the most important factor influencing the total carbon footprint of live performances. We could make a few estimations and draw some inferences on the factors involved in the emissions based on the given data set on emission figures, after performing data cleaning and data visualization using Python data visualization. The amount of electricity consumed, and its carbon equivalent are thought to have the greatest impact on total emissions and are also directly related to it. Using R Studio-Version 1.4.1717, a multiple linear regression model was built to estimate the electricity component of emissions based on the factors that cause consumptions to increase or decrease as they change. These are taken as the predictor variables and conclusions are drawn based on literature calculations as well as the linear model derived using the data provided. We have also conducted research on the travel emissions involved, as the client, Civic Digits, intends to tour multiple locations throughout Scotland to present their shows to live and digital audiences for their future post-pandemic proposal.

2.1 Introduction

Nowadays global climate change is becoming a topic of interest for the large part of the world. Humans are increasingly influencing the climate and the temperature of the Earth by burning fossil fuels, destroying forests, and raising livestock (European Commission, 2021). The term "global climate change" refers to the rise in global temperature caused by human factors. The earth surface naturally heats up by absorbing the heat from the sun, however due to presence of certain greenhouse gases(GHG) such as carbon dioxide (CO_2) and methane (CH_4) in our atmosphere the heat cannot escape the surface of the earth, eventually making the temperature even higher (European Commission, 2021). This effect is also called the greenhouse effect. Increase in the GHG in the atmosphere makes the earth warmer and warmer everyday causing more global warming. Utilisation of low-carbon energy sources and improving energy efficiency using renewable sources of energy will result in lower GHG emissions (Grce, 2021). Alisa Singer, an artist at the Environmental Graphiti® through the painting "Changing" expressed that, *"As we witness our planet transforming around us we watch, listen, measure… respond"*(MassonDelmotte et al., 2021). Organisations with the help of Creative Carbon Scotland have been aided to measure and report their carbon emissions through a survey conducted that was filled by the green champions of each of the responding companies.

The Paris Agreement under the United Nations Framework Convention on Climate Change in 2015, a legally binding international climate change treaty, was adopted by 196 parties on December 12, 2015, at COP 21 in Paris, and went into effect on November 4, 2016 ("The Paris Agreement | UNFCCC," 2021). The ultimate goal of this agreement was to keep the global temperature rise well below 2°C of warming, preferably 1.5, when compared with pre industrial levels (Versteijlen et al., 2017). The decade from 2011 to 2020 was the warmest on record, with the global average temperature reaching 1.1°C above pre-industrial levels in 2019. Global warming caused by humans is currently increasing at a rate of 0.2°C per decade (European Commission, 2021). We can measure the warming effect of these greenhouse gases using a single unit. The term "carbon dioxide equivalent," or "CO_2e," is a more accurate way of referring to different greenhouse gases in a common measuring unit, which is the amount of global warming caused by a given amount of GHG at a given point in time, expressed in terms of the amount of CO_2 that would have the same instantaneous warming effect. For example, if 1kg of methane is emitted, the equivalent amount of CO_2e is 25kg (1kg CH4 * 25 = 25kg CO2e) (Brander, 2012) .

Even before a global pandemic confined millions to their homes, advances in internet connections and service offerings had resulted in an exponential increase in the use of streaming video globally (Kamiya, 2020b). Information and communication technology (ICT) sector is considered to aid the work from home life to a very large extent. The UNDP's Global Goal for Sustainable Development can be reached with the increased access of recent technologies to large number of people (Ericsson, 2020). A large part of the sector's carbon footprint can be linked back to electricity consumption, but many key ICT players invest in renewable energy, such as solar and wind power, in a bid to lower their carbon emissions. The emissions during use emerge almost entirely from electricity consumption, but the other life cycle stages consume electricity too, for instance in manufacturing. If the ICT industry and its users only consumed electricity produced by renewable energy sources, more than 80 percent of ICT’s carbon footprint could be reduced. The three main components of the ICT sector are said to be user devices, networks, and data centres (Ericsson, 2020). The ICT sector's total life cycle carbon footprint is approximately 730 million tonnes CO_2 equivalent (Mt CO_2-eq), or 1.4% of total global greenhouse gas emissions(Malmodin and Lundén, 2018). This includes not only the total amount of electricity consumed by all system equipment while in use, but all other aspects of the life cycle also, such as the production of networks, data centres, phones, personal computers/laptops, and other user equipment (example, home routers). Furthermore, these figures include the cost of building ICT-related structures, as well as employee travel and transportation (Ericsson, 2020). Energy consumption is measured in unit called kilowatt Hour (kWh) which gives the amount of energy equal to one kilowatt of power sustained in an hour. Also, the factor used to convert kWh to kgCO_2e depends on how the power that is supplied to the national grid is generated and that varies from country to country. From the information provided by the Creative Carbon Scotland, the UK conversion factor has been declining as more low-carbon renewables have come online (it was nearly 5 just a few years ago), so it is quite low in international terms. Although not all of the electricity used in the streaming process will be generated in the UK, we do not know enough to assume a more complicated situation. For example, in the year 2020, the average of carbon emissions per hour of video streaming in Europe is estimated to be 56g CO_2e/hour video streaming. At the same time, it was 48g CO_2e/hour for UK. (Stephens et al., 2021)

2.1.1 Carbon footprint and digital performances

The term "carbon footprint" refers to a well-known concept that is widely accepted as the life cycle carbon equivalent emissions and effects associated with a product or service. (Malmodin and Lundén, 2018) We usually associate the term "carbon emissions" with traffic congestion, factory emissions, private jets and other transport, and in major metropolitan areas. Many people are unaware that, depending on the technology used, high-definition video and game streaming can result in significant greenhouse gas (GHG) emissions (“The Carbon Footprint Of ... Video Streaming and Games? It’s A Thing | Nasdaq,” 2021). Even so, as technology and digitisation increase in almost every segment of market, there is carbon emission associated by the energy consume by devices, data transfer and internet in production and use, which is all summed up under the term 'digital carbon footprint.' Digital technologies now account for 4% of greenhouse gas emissions and their energy consumption is increasing at a rate of 9% per year (Geist, 2019). The ICT sector footprint includes emissions from ICT-related services such as gaming, social media, and online advertising. “There was a lot of misinformation and misunderstanding about the carbon impact of video streaming,” said Andie Stephens, the white paper's lead author and associate director at the Carbon Trust. (Stephens et al., 2021)

2.1.2 Streaming a Video

Streaming video is also becoming increasingly popular around the world. With the onset of the pandemic, online video streaming has been considered as a great alternative to live performances and art shows throughout the world. From a life-cycle perspective, video streaming is tied to energy consumption and carbon emissions from network infrastructure. The network infrastructure includes two major sectors which contribute to the energy consumption. The sector of information and communications technologies (ICT) is divided into three sub-components: data centres that store and process data, networks (both mobile and fixed), and end-user devices (excluding devices included in the E&M boundary). (Stephens et al., 2021) All electronic equipment used for media and entertainment, such as televisions, cameras, and other E&M consumer electronics, as well as physical paper media and printing, are included in the Entertainment and Media (E&M) sector. Cinemas, theatres, and other physical site events (e.g., sports) that require physical human presence have been excluded from the E&M sector, as the focus of this study is on the use of electronic equipment for entertainment and media purposes (Malmodin and Lundén, 2018), as shown in Figure 2.

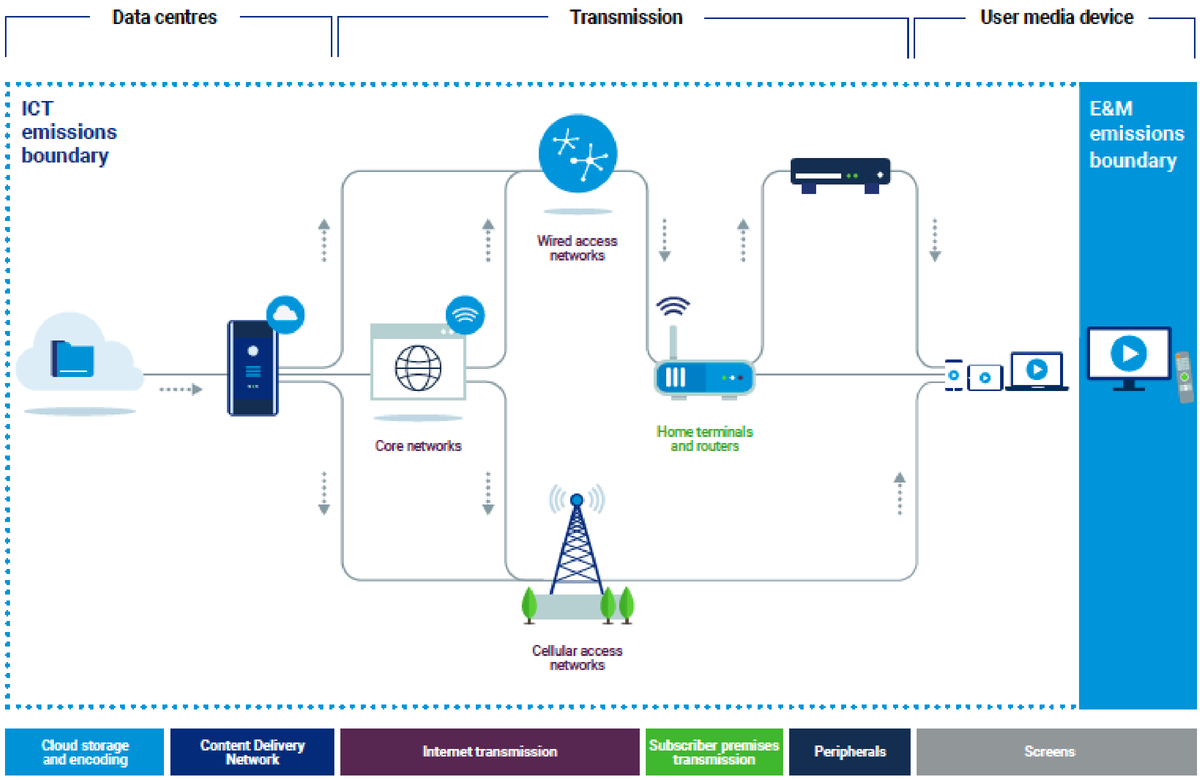

Figure 2: A process map showing the emission boundaries of ICT and E&M sectors

ICT emissions boundary – This includes originating point of video content and storage- Data Centers, Content Deliver Network and the Cloud-storage and encoding devices, the networks (both fixed and mobile) and the end-user devices like home routers and terminals which are not a part of E&M boundary.

E&M emissions boundary – This comprises of all electronic equipment used for media and entertainment purposes like TV, camera, print media, etc.. Source: (Stephens et al., 2021)

The energy consumption of each of the internet is dependent on the resolution of the user screen hence would vary with the type of device used to stream the video. We have thus considered 'The Big Data Show' by Civic Digits as our point of research and made all analysis and calculations in our report based on the same.

2.2 Aim and Objective:

The purpose of this paper is to evaluate face-to-face performances as a functional unit in order to define the relevant factors of power consumption and environmental impact of live performance production and compare it to other forms of digital performance such as zoom lectures, digital shows, video conferences and meetings, and so on. This project presents a comprehensive study to evaluate the actual energy consumption which contributes the highest towards the carbon emissions in a streamed video performance as compare to a live theatre performance. The concept of carbon footprint extends beyond the use of electricity by products. It includes greenhouse gas emissions from energy and materials used throughout a product's life cycle. The operating and lifespan energy consumption of end users, video streaming equipment, and network infrastructure (both fixed and mobile) are all part of the research. The energy consumed by audiences to attend these events is then compared, including direct fuel consumption of vehicles, public transportation (local and international), theatre or art center infrastructure, and number of travelers. We also examine how varying travel distance and video streaming duration affect the overall carbon savings generated by digital performances.

Thus, keeping these basic things in mind and further doing some literature research we will scope out the irrelevant or non-impacting features from the data and build a regression model with necessary assumptions which can help us calculate the factors on which the most important factor of emissions depends. It has been evident that a large portion of the carbon emissions is linked back to the electricity consumption, so a majority of the ICT industry and its users are switching to renewable energy sources. The main assessment and analysis would circle around the fact that the energy consumptions in a digital performance and subsequent carbon emissions are infinitesimally small and thus, in the future lowering the carbon footprints. Previous investigations have shown that, in the worst-case ICT industry will consume as much as 51% of the earth's electricity in 2030 if there is not enough energy efficient wireless networks and data centers established. Also, ICT's electricity consumption would be responsible for up to 23% of global greenhouse gas (GHG) emissions (Andrae and Edler, 2015).

Civic Digits produced 'The Big Data Show' and it is intended to be presented in both formats in the future so will be a key element in assessing and comparing the impact of live and digital presentations. With the help of some reports that have already done the analysis on ICT processes in general and understanding the technical details of 'The Big Data Show', an attempt is made to estimate the ICT footprint of digital event and then used the model obtained from live events to find inferences on lowering the annual emissions. Going ahead with some initial cleaning and exploratory analysis with the given data by Creative Carbon Scotland that records the emissions in live face-to-face performances in Scotland in the year 2018-19 the main focus of the project is to scope the most important factors contributing to the carbon footprint of a digital streamed performance or a live performance. Electricity consumption can be seen as an important factor which in turn depends on the size of the organisation i.e., the number of employees and members of staff in that company, the given office floor area which is dependent on the type of the organisation.

2.3 Literature Review

With the outbreak of the pandemic, every sector is moving to the online and digital functions. From videoconferencing in IT industries to online classes in the education sector everyone has now adopted the digital platform and ICT has taken a maximum boom in the past year since the outbreak of the pandemic and the world being forced to be under social distancing and lockdown in picture, and for the businesses to keep flourishing in the meantime. In its recommendation L.1470, the ITU Telecommunication Standardization Sector (ITU-T) estimated the ICT sector's carbon footprint for 2015 at 740 MtCO2 e, including embodied emissions. This totals to about 1.3% of the total global greenhouse gas (GHG) emissions, which is nearly five times less than the global footprint of the iron and steel sector, and it is smaller than the global footprint of many other medium and large-scale industries (ITU, 2020). There have been various research and reports which have analysed the footprint of ICT processes in general. We could adapt all the processes and then use the information on actual technical elements involved in the production specific to the "The Big Data Show" by Civic Digits, a flourishing digital theatre company, to get a better idea of the digital carbon footprint for theatre companies if they were to switch to completely online delivery mode. Civic Digits also have an ambition to present a blended programme of both live and streamed performances and they are currently investigating what this will involve.

Creation Theatre Company is a producing theatre based in Oxford, England. In its report entitled Digital Theatre: A Route to Sustainability, Creation Theatre has revealed that as compared to the pre-pandemic in-person work, the digital productions output has reduced its carbon emissions by 98% (Masso, 2021). It discovered that the greatest environmental benefit of digital theatre is reduced audience travel. Other factors compared by Creation Theatre in their report included the average distances covered by the audience that eliminated completely after the digital audience (Masso, 2021). The power usage of the show lighting, sound per person and an entire performance run is compared with the household energy consumed by its performers, technicians and presenters. The physical materials used to build up the set for one show is also considered as a comparison which also becomes zero due to digitisation. All of this helped them cut down the carbon emissions to a very large extent. Lucy Askew, Creation Theatre chief executive and creative producer in an interview said "Our digital work over the past 15 months has shown us a route to making far more meaningful carbon reductions in our industry" (Masso, 2021). According to the report, if an average-sized tour of regional playhouses with a total estimated audience of 256,000 were to access the work digitally, 197 tonnes of carbon would be saved (Masso, 2021).

Ericsson has spent over 20 years researching digital carbon emissions, amassing massive data sets and publishing findings in peer-reviewed scientific journals and at conferences (Ericsson, 2020). After comparing and analysing various published results and viewing multiple statements about carbon emissions they have shared their findings in their report- ICT and climate in 2020. They have analysed how ICT impacts climate by the carbon emissions involved, directly or indirectly, in its manufacturing, use and disposal. Air travel, excluding air transport and military flying, accounts for approximately 80% of the total aviation sector footprint (Ericsson, 2020). In 2015, the aviation industry's emissions from fuel combustion were estimated to be around 800 million tonnes of CO_2 (Miller and Façanha, 2014) whereas, the ICT sector's total life cycle carbon footprint is estimated to be 730 million tonnes CO_2

equivalent (million tonnes CO_2-eq), or 1.4 percent of total global greenhouse gas emissions (Malmodin and Lundén, 2018). From a life-cycle perspective, the actual figures from aeroplane manufacturing, airport operation, including ground vehicles, end-of-life treatment, and an accurate estimate of any other greenhouse gases (GHG) produced by the sector, including high-altitude effects, should have been taken into consideration for a balance comparison with the total cycle of ICT. However, not in terms of its scope but the similarity of its numbers, they consider just the fuel factor of emissions from an aviation industry for the comparison because of less availability of its other life-cycle components (Ericsson, 2020). After exploring more on what the ICT sector can do in order to reduce its footprint they summarised that 80% of the ICT's total carbon emissions can be reduced if they could use renewable energy sources to produce all of its electricity (Ericsson, 2020).

A lot of similarities can be drawn between the energy and information technology (IT) industries since the two industries depend largely on the services provided by one another (Costenaro et al., 2012). Businesses and individuals are flocking to cloud-based software, offsite data warehouses, and streaming media services, eager to free themselves from the burdens of owning and maintaining IT equipment and infrastructure (Costenaro et al., 2012). Data transfer is one important aspect to keep in mind when thinking of digital performance streaming. Another research in the same scope entitled *'The Megawatts behind Your Megabytes: Going from Data-Center to Desktop*' answers this question of what is the energy consumption of data transmission and the equipment that it uses. (Costenaro et al., 2012) The infrastructure is generally configured in the manner depicted in Figure 3 below, with three tiers of increasingly advanced points of presence (POPs) to transmit packets between end-user devices.

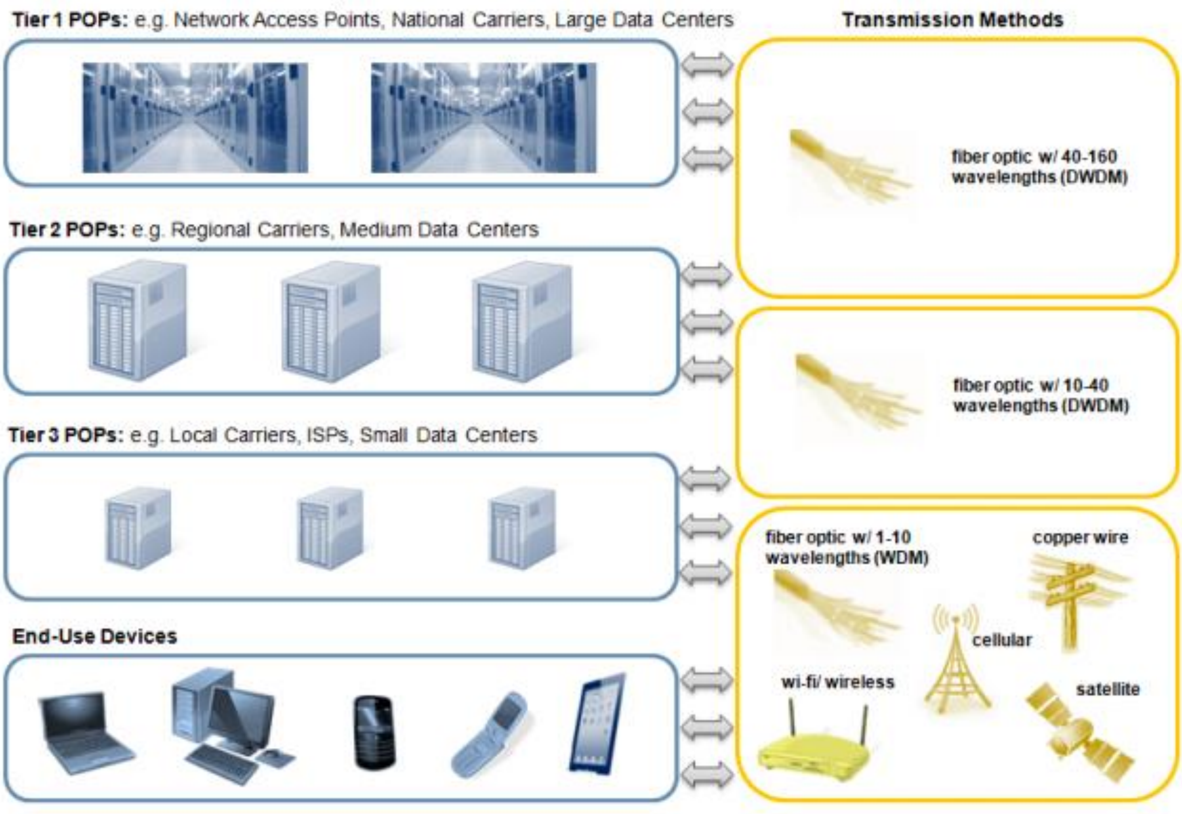

Figure 3: Infrastructure of the Internet at different points of presence (POPs) and the Transmission Methods used to transmit packets between end-user devices. Source: (Costenaro et al., 2012)

According to Raghavan and Ma's estimate, the internet's wall socket power ranged between 84 and 143 GW (Raghavan and Ma, 2011). Costenaro and Duer in their research however refined the estimate by further dividing the equipment and activities into tiers as described in Figure 3 where network access points, national carriers and large data centers are considered Tier 1 POPs, smaller regional carriers and medium-sized data centers are referred to as Tier 2 POPs and the even smaller data centers, local carriers and our internet service providers (ISPs) are said to be Tier 3 POPs (Costenaro et al., 2012). As obvious as it may sound, tier 1 devices will draw the highest energy and gradually reducing till the end-user devices which are smartphones, laptops, tablets, etc. The results from their analysis show that the total average of energy of internet and equipment used is 141GW (Costenaro et al., 2012). The sub-division of the tiers is given by the following chart. If the Internet's global average energy draw is 141 GW, the analysis suggests, the system will require well over 100 base load power plants. Using the segmented analysis in Figure 4, where the blue portion represents the end-user's share of consumption, the red color in the plot is symbolic of the share of energy used in transportation of data which is reduced by use of data centers and the green color is depictive of the portion of energy consumed by the data centers or the POP (Costenaro et al., 2012). The figures estimates that only 38% of that load is associated with end user devices, 14% with transmission and communication aspects, and 48% with network resources in data centers and server rooms (Costenaro et al., 2012).

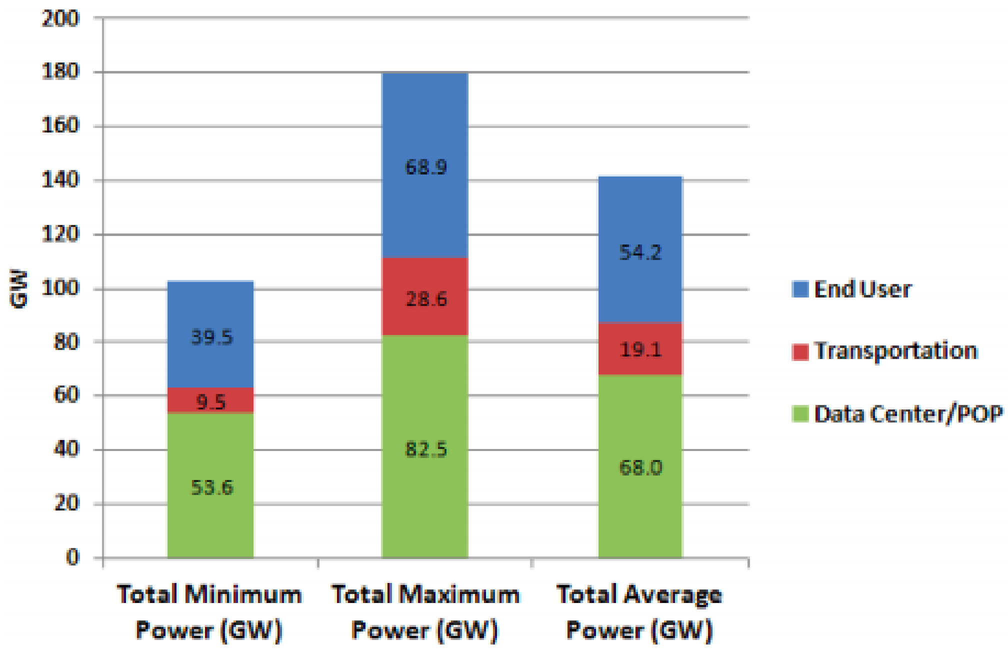

Figure 4: Total Internet Energy Usage (GW) based on the segmented analysis for End User (blue), Transportation(red) and Data Centre/POP(green). Source: (Costenaro et al., 2012)

Growing technological trends are also making a collective impact. Internet-connected mobile devices, such as tablets and smartphones, are shrinking in size and becoming far more energy efficient in a variety of ways. One of the most noticeable changes across most devices is the transition from CRT to LCD screens, the latter of which uses less electricity (Grce, 2021). As published in the report by Sergio Grce, CEO of iSIZE, the overall footprint of streaming video is heavily influenced by how electricity is generated. The internet is now used by approximately 4.1 billion people, or 53.6 % of the world's population (Grce, 2021).

Individually, streaming video over fiber optic cables emits the least amount of CO_2, at a rate of 2grams(g) per hour. On the other hand, streaming over next-generation mobile technology, also known as 5G, would result in 5g of CO_2 emissions per hour (Grce, 2021). Taking into consideration that the average use of 3G mobile technology emits 90g of CO_2 per hour.

Transportation contributes a truck load of share in the global carbon emissions. Found in the research paper by Robert Sausen & Ulrich Schumann on climate change, aircraft emissions can affect the climate in several ways: (1) they emit substances that are radiatively active (e.g., CO_2); (2) they emit substances that produce or destroy radiatively active substances (e.g., NO_x, which modifies O_3 concentration); and (3) emissions cause the formation of new clouds (e.g., contrails) (Sausen and Schumann, 2000). Until 1995, aviation CO_2 emissions increased atmospheric CO_2 concentration by 1.4 ppmv (1.7 percent of the anthropogenic CO_2 increase since 1800). By 1995, the global mean surface temperature had risen by approximately 0.004 K, and sea level had risen by approximately 0.045cm (Sausen and Schumann, 2000). However, in case of our research we will just be looking at the CO_2 fuel emissions only and keeping other components aside due to the scope of data availability. The Department for Environment, Food, and Rural Affairs (DEFRA) of the United Kingdom also publishes data on the energy and carbon emissions of various modes of transportation. (Belenky, 2011). DEFRA only included direct emissions and the lifecycle impact of fuel from extraction, refinement, storage, and transportation. Ong, Moors and Sivaraman have compared these to the emission figures fairly based on a research by Lenzen who considers all life-cycle components such as direct energy consumption from fuel and electricity, as well as indirect emissions from fuel production, vehicle lifecycle, electricity generation, maintenance costs, and related infrastructure construction (roads, railways, stations, airports etc.) (Lenzen, 1999).

There have been various agreements and disagreements in the emissions of streaming a video. George Kamiya in his report on 11 December 2020, a version of which was also published in Carbon Brief as well, exposed the flaws in a widely reported estimate of the emissions from 30 minutes of Netflix watching (Kamiya, 2020a). Data centre and data transmission network energy intensity figures have been updated to reflect more recent data and research. As a result, the central IEA estimates of the emissions for one hour of streaming video in 2019 were 36 gCO_2e, down from 82 gCO_2e in the original February 2020 analysis (Kamiya, 2020a). A number of recent media articles from the United States have iterated the claim that “emissions from watching 30 minutes of Netflix [1.6 kg of CO_2] is the same as driving almost 4 miles” (Kamiya, 2020b). The figures come from a report in 2019 on the unsustainable and growing impact of online videos by a French think tank, Shift Project who later also updated their results in a follow up article in 2020 to correct a bit/byte conversion error, revising the original "1.6kg CO_2e per half hour" quote by 8 times to 0.2kg CO_2e per half hour (Geist, 2019). From the Cisco forecasts, the share of streaming energy consumed by used devices, transmission and data centres for different devices is shown in a graph (Kamiya, 2020a). The graph (Figure 5) shows the share of streaming energy from devices, data transmission and data centre and comparing it with the figures from the two Shift Project reports in 2019 and 2020, the figures from IEA reports and the results for individual streaming devices as estimated by George Kamiya (Kamiya, 2020a).

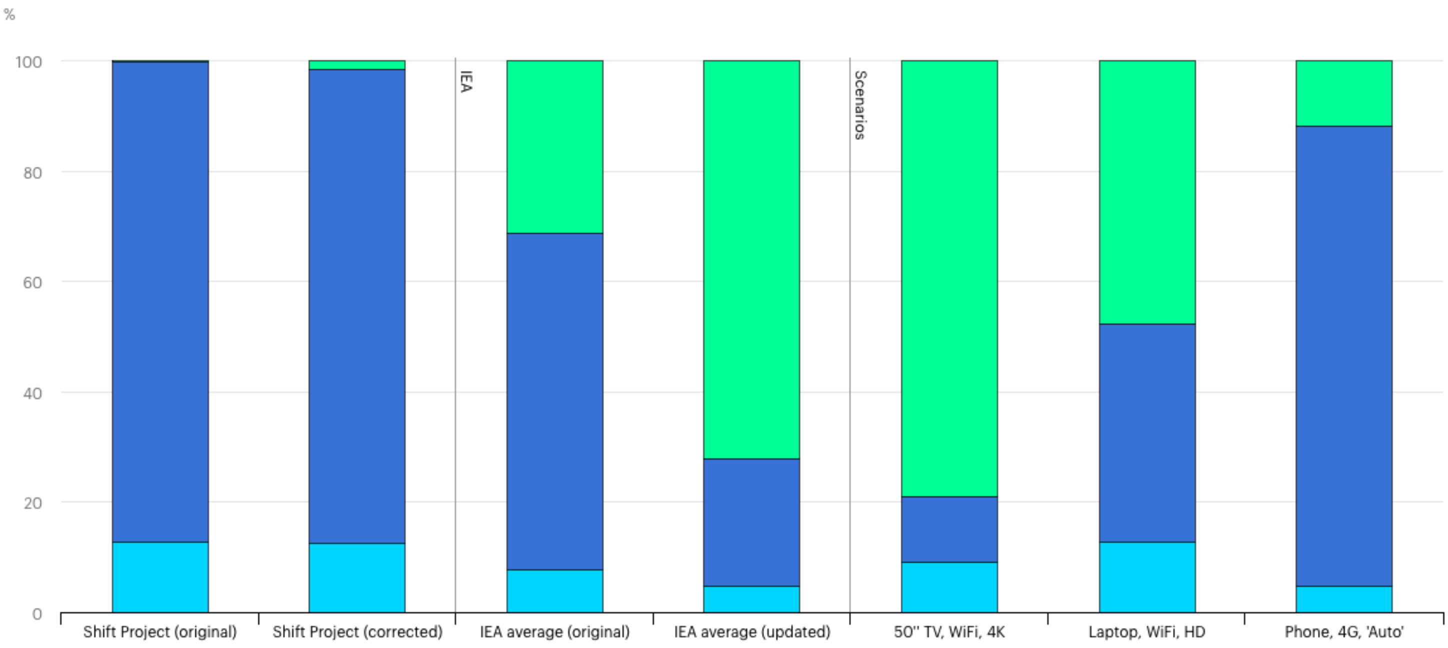

Figure 5: Illustrating the share of streaming energy use from devices(light green), data transmission(darker blue) and data centers (light blue) (Kamiya, 2020a)

A similar case could be studied in a report presented in an article in the Netherlands, that explored carbon emissions associated with student (and staff) travel at several Dutch HEIs. That research included an evaluation of HEIs' (Higher Education Institutions) reported GHG emissions as well as the findings of in-depth interviews with HEI professionals. It also provided findings on the pros and cons of online education based on an analysis of interviews with HEI professionals (Versteijlen et al., 2017). In the literature review of the study, they found that there are three main analysis categories which form the total footprint for the consumption of energy by an HEI for on-site as well as off-site emissions i.e., building energy, travel and procurement (Ozawa-Meida et al., 2013). With the online teaching in place due to the COVID-19 pandemic, we can take relevant points from this research as to compare online and offline performances in our research. A basic 3-step approach is used to estimate carbon footprints in all analysis categories. First they determine the activity/consumption in each of the sectors in terms of kWh energy used in building, kilometres travelled by the student and staff and the money spent on procurement. Then they have derived the respective GHG emission factors in terms of kg e per kWh, kilometre(km) or pounds spent. Finally, multiplying the two to obtain the overall carbon footprint in kgCO_2e for each of the sectors (Ozawa-Meida et al., 2013). The emission factors used in this study are based on the Defra/DECC Guidelines for Greenhouse Gas Conversion Factors for Company Reporting (Wiedmann and Barrett, 2011). Similarly, we can calculate the energy in production and operation of all the associated devices, data centre and internet infrastructure used in a video streaming and multiply it with the carbon emission factors given by the Scottish government to finally derive the carbon footprint of theatre organisations switching to the digital world. Results suggested that in De Montfort University in 2008-09, the total consumption based GHG emissions were nearly 51,080 t CO_2e, in which energy use accounted for 34% of total emissions, transportation accounted for 29%, and procurement accounted for 38% (Ozawa-Meida et al., 2013). The staff and domestic student's commute emitted 18% of the total GHG emissions, however the international students alone produced 6% of the total emissions due to travel (Ozawa-Meida et al., 2013).

While reviewing how digital streaming is a better and greener alternative that aids in the substitution of physical processes for virtual ones in order to significantly reduce emissions, we also came across a paper by Ong, Moors and Sivaraman that evaluates the total energy, carbon and time costs of videoconferencing as compared to in-person meetings which included network and videoconferencing equipment operating costs, equipment lifecycle assessment, and meeting participants' time costs (Ong et al., 2014). We could easily make use of this since videoconferencing can be a good proxy for digital streaming of performances, whereas the in-person meetings could be related to the live performances attended by audiences. The Life-Cycle Analysis of network and videoconferencing terminal equipment is thoroughly explained, and the carbon emissions calculated respectively (Ong et al., 2014). Some of the calculations we have performed in further stages of the research were directly influenced by this report. They have calculated the total power as a sum of operational energy as well as the embodied energy consumption for the individual components. The two equations used are given as equation (1) for operational energy and equation (2) for embodied energy (Ong et al., 2014).
For overall operating power P_{op} calculation, they have used equation

$$P_{op} = \sum_{i=device} n_i k_i P_i + I_{NW(op)} R, \tag{1}$$

where $n_i$represents the number of devices of type *i,* P_i represents their individual power consumption, $I_{NW(op)}$ is the operational network energy intensity, *R* is the average rate of data or bandwidth of the conference and k_i is the use factor, i.e. the amount of the device's use dedicated to the videoconference. Similarly, for embodied power P_{em} calculation, equation (2) is used (Ong et al., 2014).

$$P_{em} = \sum_{i=device} n_i k_i \frac{E_i}{U_i} + I_{NW(em)} R, \tag{2}$$

where, $n_{i,}$ $k_i$and *R* as same as above and E_i represents the individual embodied energy of device type *i,* U_i is the total number of hours the device operated over its lifetime and $I_{NW(em)}$ is the embodied network intensity (Ong et al., 2014). The above-mentioned equations are then applied to two different configurations of teleconference to estimate the upper and lower bounds of power consumption and hence, the carbon emissions. We do not get into so much of details as of the LCA and individual device configurations, but we can adapt parts of this paper for our calculations and analyses. However, analysis categories may differ depending on the purpose of study and the impacts being examined.

2.3.1. Estimating the Energy Consumption in Data transmission of streaming a video:

To put it in context, the estimate for the average carbon footprint of a half-hour Netflix show is equivalent to driving around 100 metres in a conventional car (Kamiya, 2020b). The electricity consumption is dependent on the internet network transmission, the data centers used and the end-user devices and video streaming peripherals. We will elaborate more on each of these moving forward.

Network and Transmission (Core and Access) :

The metric for average energy per data volume [kWh/GB] (kilowatt hour per gigabyte of data transmitted) is used to allocate internet network electricity (Stephens et al., 2021)d. According to estimates, approximately 3.97 billion people worldwide had Internet access in 2019. By mid-2020, Asia was projected to have the most Internet users (2.5 billion), while Europe can have 728 million (Statista Research Department, 2020). As a result, it is to blame for both the ever-increasing global demand for energy and the ever-increasing CO_2 emissions. We compute the total energy consumed by the Internet operations and devices that make up the internet and summarise our findings. From the average calculated in 2019 for data transmission through a Wi-Fi the electricity consumed is 0.018 kWh per viewing hour (Kamiya, 2020b). 'The Big Data Show' is a 70 minute online streaming show including performances and interval where games are played, so let us take 1 hour of streaming into consideration. From the conventional approach used in the Carbon Trust study, the boundary is limited to that shown in Figure 2. Thus, the results assume the Fixed network intensity as 0.0065kWh/GB and for a Mobile network to be 0.1kWh/GB (Stephens et al., 2021). For a home router it is considered as 0.025kWh/GB which is applicable for only fixed network viewing. With data transmission rates derived from (Netflix, 2021) for a fixed network and a standard definition it transmits data at 1GB/hr. Full definition video is streamed at 3GB/hr and Ultra definition 4K is streamed at 7GB/hr. For a mobile network it is 0.17 GB/hr, 0.25 GB/hr and 3GB/hr respectively. (Netflix, 2021)

User Devices:

The largest portion of the sector's overall carbon footprint is accounted for by user devices (including phones, tablets, and computers). To perform the functions of a CODEC, software-based online video systems rely on the processing power of a standard PC. Raghavan and Ma have estimated the power consumption of a present day laptop and desktop personal computers (PC) to be 40 W and 150 W respectively. (Raghavan and Ma, 2011) The results highly depend on the viewing device, network connection type, and resolution used. A 50-inch LED television, for example, consumes 100 times the amount of electricity as a smartphone or 5 times as that of a laptop. However, if we consider 'The Big Data Show' by Civic Digits to keep the viewing device as a Laptop streaming HD video, the energy consumption would be close to 0.022 kWh (Kamiya, 2020b) and considering the same show of 70 minutes streaming would give us around 0.026kWh of electricity. Multiplying this by the carbon emission factor for electricity 0.26 per kWh in Scotland, ("Some-quick-carbon-conversions-2019-20," 2019) we get the CO_2 equivalent for 1 laptop as ~0.0068 $kgCO_2e$. Similarly, calculating for 1 4Gsmartphone for using the application we can get 0.00234

$kgCO_2e$(Kamiya, 2020b). If we consider, an audience count of 500-1000 people as it was estimated by Civic Digits member Rupert Goodwins, we can calculate the total $kgCO_2e$ emitted by user devices for a single show, including all laptops and smartphones, to be 4.57 – 9.14 $kgCO_2e$.

Data Centres and CDN :

In 2019, global data centre electricity demand was around 200 TWh, or about 0.8 percent of total global electricity demand.(IEA, 2020) Data transmission network technologies are also rapidly becoming more efficient. In 2020, electricity consumption of data networks is expected to rise to around 270 TWh , based on current efficiency improvement trends (IEA, 2020). A data centre houses both cloud storage and encoding services (Stephens et al., 2021). A CDN (content delivery network) is a platform of servers that are highly distributed, which help in minimizing the delays in loading a video or webpage content on your device screen by reducing the geographical distance between end-user and the streaming server. User can thus, enjoy same high-quality content without buffering times. CDNs operate by keeping a copy of static files on each of their servers around the world. This means that when a user visits your website, these files are served from a local CDN server, reducing the volume of data that must travel. Reducing data transfer is also better for the environment! (Butterworth, 2020) Demand for data centre and network services will continue to rise sharply, fuelled in part by the rapid growth of streaming video and gaming. Between 2019 and 2022, internet video traffic is expected to more than double to 2.9 ZB, while online gaming traffic is expected to quadruple to 180 EB.(Cisco, 2019) Working with electricity utilities, regulators, and project developers, data centre operators investing in renewable energy should seek to identify projects that maximise benefits to the local grid while also reducing overall GHG emissions. Additional efforts by companies and regulators could maximise system-wide benefits and reduce emissions.(IEA, 2020) Carbon trust's conventional approach has derived the energy intensity of 1.3Wh/hr in 2020 from a selection committee of DIMPAC. (Stephens et al., 2021)

CODECs:

Codec choice and engineering is a very big part of online video. Power consumption of almost 80 W for high-end CODECs is estimated, which represents the average power consumption among top-of-the-line CODECs from the major manufacturers (Polycom HDX 9000, LifeSize Room 220 and TANDBERG C90) (Constable, 2011) Similarly, the average active power consumption for entry-level CODECs tested in (Constable, 2011) (LifeSize, PassPort and TANDBERG C20) is 26 W. The LCA carbon emissions are calculated to be 187* kgCO2 e for high-end and 61* kgCO2e for entry level CODECs. However, as per Scotland's carbon factor it would be even reduced to 81 kgCO2 e to 26.4 kgCO2e.

*Calculated using 0.6 as emission factor.

Video Peripherals:

Although peripherals have a low power consumption, we include them here for completeness. Cameras are estimated to consume 9.5 W (Constable, 2011) which accounts to 0.00247 kg CO_2e, while sound systems consume 4.1 W . (Ong et al., 2014) which leads to as low as 0.001 $kgCO_2e$. Based on the operating power of a studio quality microphone, we

estimate a power consumption of 2.5 W (0.000065 kgCO2 e) for a studio quality microphone and even less for lower quality radio and wireless microphones. (Ong et al., 2014)
All of these are so infinitesimally less compare to other things that it can all be ignored.

2.3.2. Emissions by travel:

Outdoor air pollution caused 3.2 million premature deaths worldwide in 2010 and is one of the top ten health risks (Miller and Façanha, 2014). Motorized transportation significantly contributes to outdoor air pollution, especially near major highways and in densely populated urban areas. The contribution of aviation to climate change – 3.5 percent of warming or 2.5 percent of carbon dioxide emissions – is frequently underestimated. In comparison to other sectors, it currently accounts for a small portion of emissions. (Ritchie, 2020) According to analysis by DEFRA if we only consider the fuel emissions for our study we get that the CO_2 equivalents in $kgCO_2e$ per kilometre of journey for various transportations. From the two studies by Lenzen and DEFRA, we can compare the travel emission figures in Table 1..

Table 1: WHOLE LIFECYCLE (INC. OPERATING) COSTS (LCA) OF TRANSPORT (Ong et al., 2014)

	LENZEN (Lenzen, 1999)	DEFRA(Belenky, 2011)
Mode of transport	Entire Life-cycle emissions ($kgCO_2e$/per km)	Fuel emissions only ($kgCO_2e$/per km)
Flight (International)	0.25	0.13
Flight (Domestic)	0.49	0.20
Train	0.17	0.06
Private Car	0.34	0.24

Therefore, as far as travel is eliminated from the digital switch to performances and electricity being a major contributor to the carbon dioxide production. So, a majority of the ICT industry and its users are switching to renewable energy sources. It is of great importance that with increase in the demand for technology, the way of production is also improvised in a way to minimize the carbon emissions.

2.4 Methodology

The data provided by Creative Carbon Scotland contains information of the annual carbon emissions of over 100 art production companies for the events over the year 2018-19 and the data is protected and not open source. It contained annual raw secondary data of companies with information on electricity and other utility consumptions by different types of organisations, the number of employees in the organisation, the floor area of each organisation as well as the travel emissions by each of the organisations during the hosting and performance of live shows over the entire year by different modes of transportation like train, bus, taxi, flights and company vehicle. The different organisations that were considered in the investigation and analysis are theatres, art centres and tenant or touring organisations all over Scotland which are given as types 1, 2 and 3 respectively in the dataset. The carbon emission factors for all the utility consumptions such as water, gas, liquid petroleum gas (LPG), fuel oil, electricity, landfill measures, biomass, dry mixed recycling and food waste on the site. were also given and using them the carbon dioxide equivalents (CO_2e) of each of the consumptions i.e., the standard unit to calculate carbon emissions and footprints, are calculated and which gives us the final processed dataset to work on. Beginning with cleaning of data, we first renamed the columns as per convenience and ease of use for further exploratory analysis. By observing the given dataset of emissions data we could figure out that only the top 137 rows contains data, so we pick those 137 organisations to further use in analysis.

Next we check the column having total annual carbon dioxide equivalent of the organisations (Total CO2e) is nothing but the sum of all the components including electricity equivalent (Elec CO2e), gas equivalent (GasCO2), water equivalent (WaterCO2), LPG equivalent (LPGCO2), Fuel oil equivalent (FueloilCO2), Biomass equivalent (BIO CO2), landfill measures equivalent (landf CO2), dry mixed recycling equivalent (DMR CO2), food waste equivalent (Food CO2) and the carbon emissions from various transportations taken by staff and audiences to attend the show. Checking for any null values in the top 137 rows, we found that there were 27 rows that do not contain any data for the CO_2 emissions for any of the above mentioned columns and also the total CO_2e of these 27 rows was also null. Thus we removed all of these rows of data as well. Completing this we again checked for blank values in any other variables we found 2 companies who did not provide information on the number of employees. Thus, looking through the company websites and finding a rough educated guess of the employee count, we replaced both of these null rows with respective values obtained.

The next step was to identify the response variable in the model, which should be the most important input factor influencing the total emissions which can be used to calculate and compare digital performance emissions. As we know, ICT's electricity consumption would be responsible for up to 23% of global greenhouse gas (GHG) emissions.(Andrae and Edler, 2015) So to make an educated comparison of the carbon footprints between traditional live audience shows and digital streamed shows we can consider 'electricity' as the major and most important factor and our dependent y variable to design the model around it. From the given data our independent x variables can be found by visualising the data, including some important factors such as number of people working, the office size which also depends on the type of the event company that we consider, might impact or influence the electricity consumption of the organisation and performance sites. Next we picked the significant features like travel and aviation that still made a lot of relevance in the carbon emission figures from the column of the Total CO2e in live performances. Initially we pick the covariates as employees, Org_type , Floor area, Elec CO2e, GasCO2 , WaterCO2 , LPGCO2 , FueloilCO2 , BIO CO2 , landf CO2 , DMR CO2 , and Food CO2 against the Total CO2e as our full model. We applied linear model on all these to find the summary in R studio. Then using appropriate selection methods, we can pick the most relevant variables in the model for the calculation of live performances. The total carbon emissions for live performance in the dataset is nothing but the sum of all the CO_2e from different utilities and travel. Thus, the linear model will be a nearly perfect fit if we ignore employees, Org_type, Floor area from the columns. The model details are present in the appendices as live performance model.

2.4.1 Simple Linear Regression :

Regression analysis is a statistical approach for determining the strength and connection between two variables. However, regression is not restricted to two variables; it may demonstrate a link between two or more variables. The regression findings aid in estimating an unknown value based on the connection between the predictive factors (Liu, 2016). When just one input variable is defined, the regression is known as Simple Linear Regression. We model a linear connection with the target variable (dependent) using a single variable (independent). We do this by fitting a model to the connection.
The four main components of basic linear regression are:

- Dependent Variable – The variable to be estimated and forecasted.
- Independent Variable – A variable that is utilised to estimate and predict.
- Slope — Line angle of the fitted regression line
- Intercept — The point at which regression line crosses the y-axis (Rawlings, 1998).

$$y = \beta_0 + \beta_1 x, \tag{3}$$

where y is the dependent variable, β_0 is the intercept, $\beta_1$is the slope that describes the change in response with unit increase in independent variable x.
So, from equation (3) the slope and intercept are the coefficients/parameters of a simple linear regression model, when we compute the regression model, and just estimates these two. Finally, attempting to identify the best-fitting line describing the data from an unlimited number of lines. To calculate the slope of a line, take a random section of the line and divide the change in x by the change in y.

When there are more than one predictive variable or covariate, the regression is known as Multiple Linear Regression. The regression model is also referred to as Ordinary Least Squares Regression when we are trying to identify the "best fit line." This simply indicates that we are employing the method with the least total of squared mistakes. The error is defined as the difference between the anticipated and actually values. The difference is squared, resulting in an absolute difference, and then added together (Rawlings, 1998). Formally, the standard model for multiple linear regression, given n observations is given by equation (4)-

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots \beta_p x_{pi} + \varepsilon_i , \quad for\ i = 1,2, \dots n, \tag{4}$$

where y_i is the dependent variable, β_0 is the intercept, β_1 to $\beta_p$are the individual coefficient of the independent variables x_{1i} to x_{pi} and p being the number of covariates in the model that are impacting the response variable. ε_i are the residual errors, i being the number of observations. The fitted values (i.e., the predicted values) are the y values that result from plugging our x values into our fitted model (Liu, 2016).

2.4.2 Multiple Linear Regression for Digital performance- Electricity consumption:

For our study and dataset we have multiple factors influencing our response variable. Thus, we use multiple linear regression. The components of multiple linear regression are similar to simple linear regression, however instead of a single X variable as independent variable that estimates the dependent variable we have multiple variables that can also be known as covariates. We could infer from all the literature studies so far that the lion's share of digital performance's footprint is linked majorly with just the electricity consumption from all the equipment, internet transmission and operations of devices and data centres, and form a model that could give us the Electricity consumption based on other factors that make most impact in the electricity used by these organisations. Multiple linear regression was carried out to investigate the relationship between the number of employees in each company, the type of organisation, the Gas utilisation and the floor area of each company. This model will be referred to as digital performance model.
A typical call may look like:
`myfunction <- lm(formula, data, …)` (Liu, 2016).
and it will return a fitted model object, here stored as myfunction. This fitted model can then be subsequently printed, summarized, or visualized; moreover, the fitted values and residuals can be extracted, and we can make predictions on new data (values of X) computed using functions such as summary(), residuals(),predict(), etc (Liu, 2016).

2.4.3 Selection of Variables:

For multiple regression, there can be a number of possible combinations and outcomes of models possible. We have method of selecting variables where we can identify the significant variables based on various selection criteria in statistics. Various techniques of feature selection are there in theory to balance between having too few variables and too less variables i.e. underfitting or overfitting the models (Tripathi, 2019). These can include non-automatic and automatic variable selection methods. One of the automatic variable selection methods we are using is called the step-wise selection. In this two nested models having a single variable difference are compared and have the same result when the extra variable coefficient from the 2nd model is made zero (Rawlings, 1998). This is to get a model with the least possible covariates that perform the same. The assessment is made by conducting hypothesis tests using the t-test or F-test statistics. One of the popular methods used in R is the step-wise selection method. It consists of iteratively adding and subtracting the model predictors. We begin with no predictors, then add the most significant predictors, and finally, after adding each new variable, we remove any variables that no longer improve model fit (CRAN, 2021). To measure the improvement of models various criteria exist in statistics like adjusted R-squared, . The criteria used in step-wise selection is Akaike information criterion (AIC) which estimates the prediction errors and the quality of the predictive models. The AIC is compared to check if it is increased or decreased by adding or subtracting a particular variable and the best model is identified with the highest possible AIC (Tripathi, 2019). We will use this in further refining the model.

2.4.4 Checking for Assumptions:

When we use linear regression (simple or multiple) to model the relationship between a response and a predictor, we make a few assumptions. These assumptions are essentially conditions that must be met before we can draw conclusions about model estimates or use a model to make a prediction.

There are 4 basic assumptions that we check for in our data:

1. The residuals are normally distributed- All variables in the linear regression analysis must be multivariate normal. This can be assessed using a Normal Q-Q plot, which can be plotted using the results from our linear model for digital performances. This is usually a scatter plot with residual values on y axis and the fitted values on the y axis. When the data is not normally distributed, a non-linear transformation (e.g., log-transformation) may be used to resolve the problem (Rawlings, 1998).

2. The residual errors are independent- Independence means absence of any kind of association between any collection of residuals. By producing pair plots and/or heat maps, we may test this notion. We can check this by plotting the residual values against the fitted values or by plotting the model residuals against the dependent variables in the model (Liu, 2016).
3. The residuals have constant variance- The variability of the dependent variable being equal across the independent variable values is referred to as homoscedasticity. We can test this assumption by plotting the model predictions and residuals on a scatterplot and searching for the residuals to equal throughout the regression line (Rawlings, 1998).

4. There is a true linear relationship - Linear regression requires a linear connection between the independent and dependent variables. Outliers must also be checked since linear regression is susceptible to outlier effects. Scatter plots are the best way to test the linearity assumption (Rawlings, 1998).

The qqplot() function generates a Normal Q-Q (quantile-quantile) plot for any distribution. If we directly plot the linear model in R, it gives 4 plots out of which Q-Q plot can be shown by the function plot(<linear_model>, which=2). Figure 6 shows an example plot of normally distributed data.

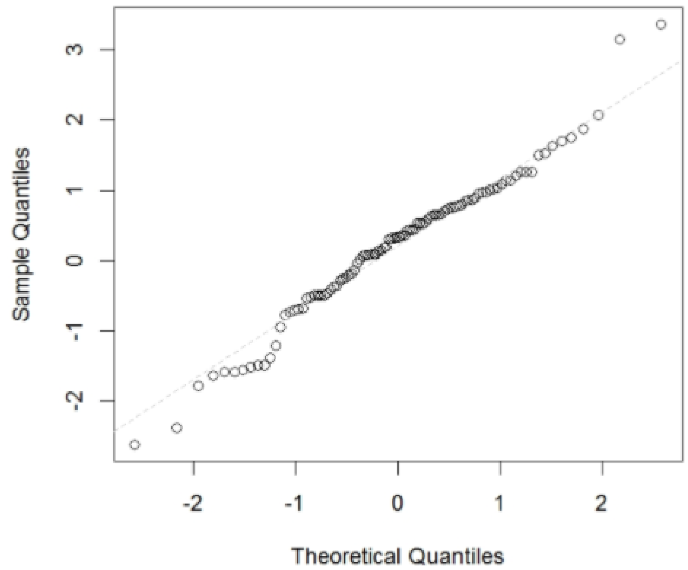

Figure 6: Example of Q-Q plot where residuals are having perfect normality

Any significant deviations from the straight line indicate that the data are not normally distributed. The plots in Figure 7 depict the possible deviations from normalcy are associated with errors having non-constant variance as indicated below(Chouldechova, 2021).

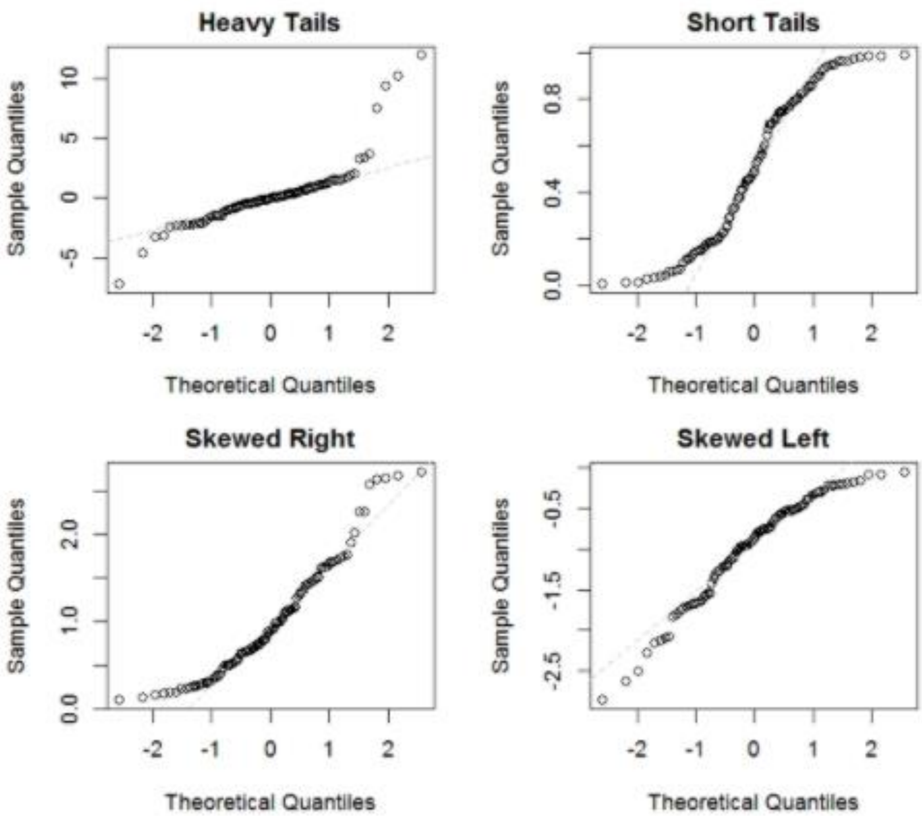

Figure 7: Different types of Q-Q plot to show the departures from normality of residual errors

In Residual vs Fitted values plot, the residual errors are plotted on the y-axis and the fitted values or estimated responses from the linear model are plotted on the x-axis. Non-linearity, uneven error variances, and outliers are detected using the plot. Here are the properties of a well-behaved residual vs. fits plot, as well as what they say about the suitability of the basic linear regression model (Rawlings, 1998):

- The residuals "bounce" about the 0 line at random. This implies that the assumption of a linear connection is acceptable.
- Around the 0 line, the residuals form a "horizontal band." This implies that the variances of the error terms are the same.
- There is no one residual that "stands out" from the fundamental random pattern of residuals. This implies that no outliers exist.

If the errors have constant variance, the spread of the residuals across the range of the fitted values will not change (Chouldechova, 2021). The function plot(<linear_model>, which=1) in R gives the residual vs fitted values plot as shown in Figure 8.

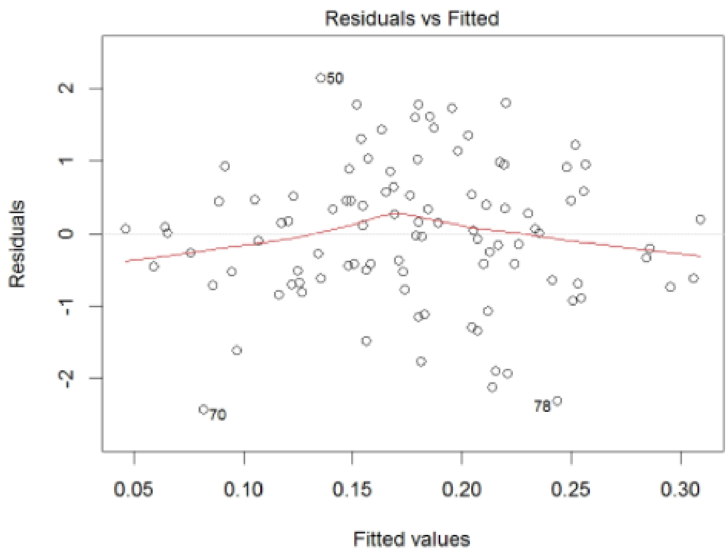

Figure 8: An Example of residuals vs fitted plot for checking the constant variance of errors assumption

Common departures from this assumption are overdispersion, which would be indicated by a fanning out of the residuals (right plot Figure 9) or under dispersion, which would be indicated by a funnelling in of the residuals (left plot Figure 9). Deviations from this assumption and the normality assumption can be reduced by transforming the response variable with the box-cox transformation family.

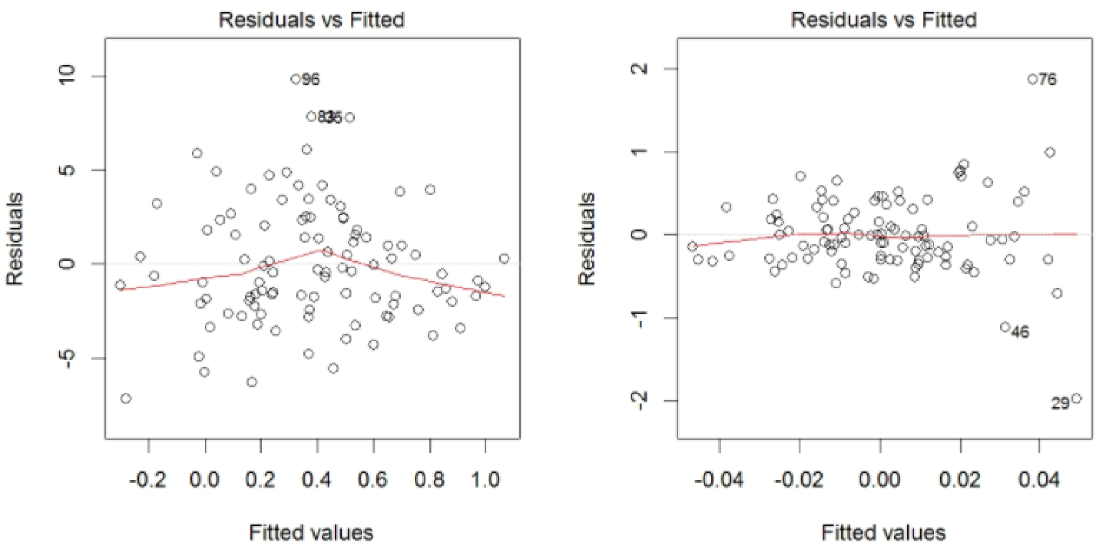

Figure 9: Example of over dispersion plot(right) and under dispersion plot (left) (Residuals vs Fitted) for checking the independence assumption

2.5 Analysis and Results

We have historic emissions data available for several types of art organisation which are touring companies, theatres and art centres. The initial cleaning involved removing the irrelevant columns and replacing null values if any with appropriate methods of mean and median comparison for replacement or educated guesses using the annual reports for information on individual companies and the shows performed, or with the help of some researches held earlier on the significant factors influencing the emissions. Performing exploratory analysis to find important and relevant data for our research using Python 3 in Jupyter Notebook is the next step and with the help of this visualisation from the given data for the year 2018-19 for the live theatre and touring companies for their offline shows we could find a few inferences as discussed further.

By plotting the total carbon emissions grouped by organisation type which is theatres, art centres and tenants we could see that the highest amount of total annual emissions were caused by theatre companies (type 1) equal to 3551474 kgCO_2e from almost 12 companies comprising of 382 employees and for the tenant or touring companies i.e. given to be as (type 3) is somewhere around 2518152 kgCO_2e from data, for a total of 71 organisations and 447 workers.

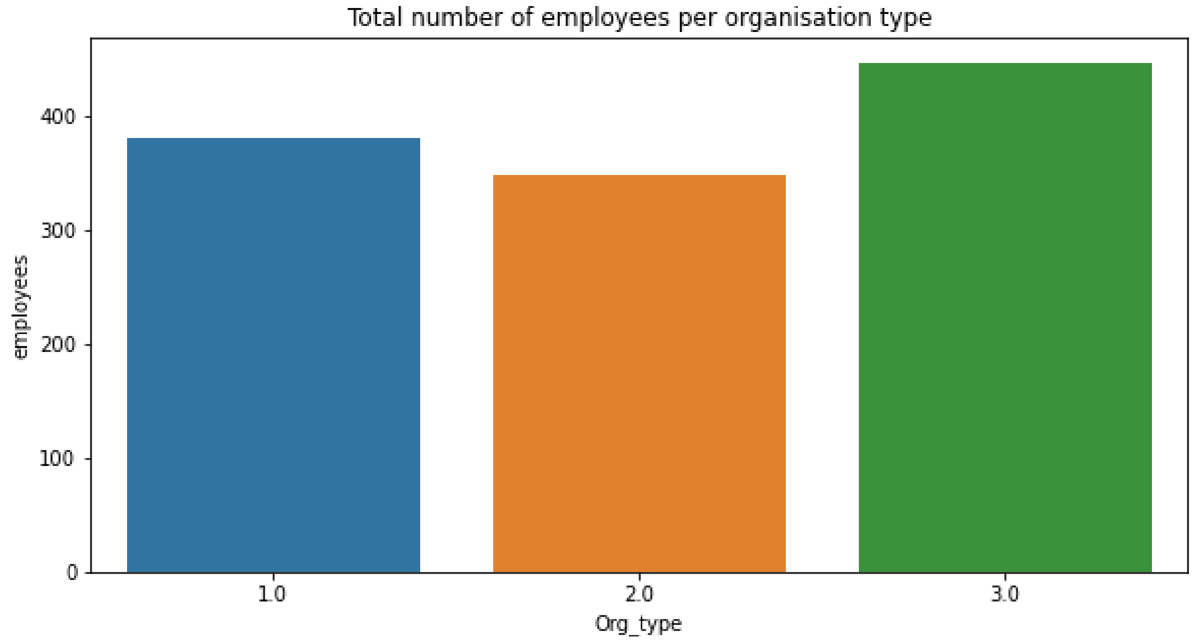

Figure 10: The plot of number of employees and staff in the year 2018-19 working with each organisation type- theatre companies (blue), art centers (orange) and touring organisations(green)

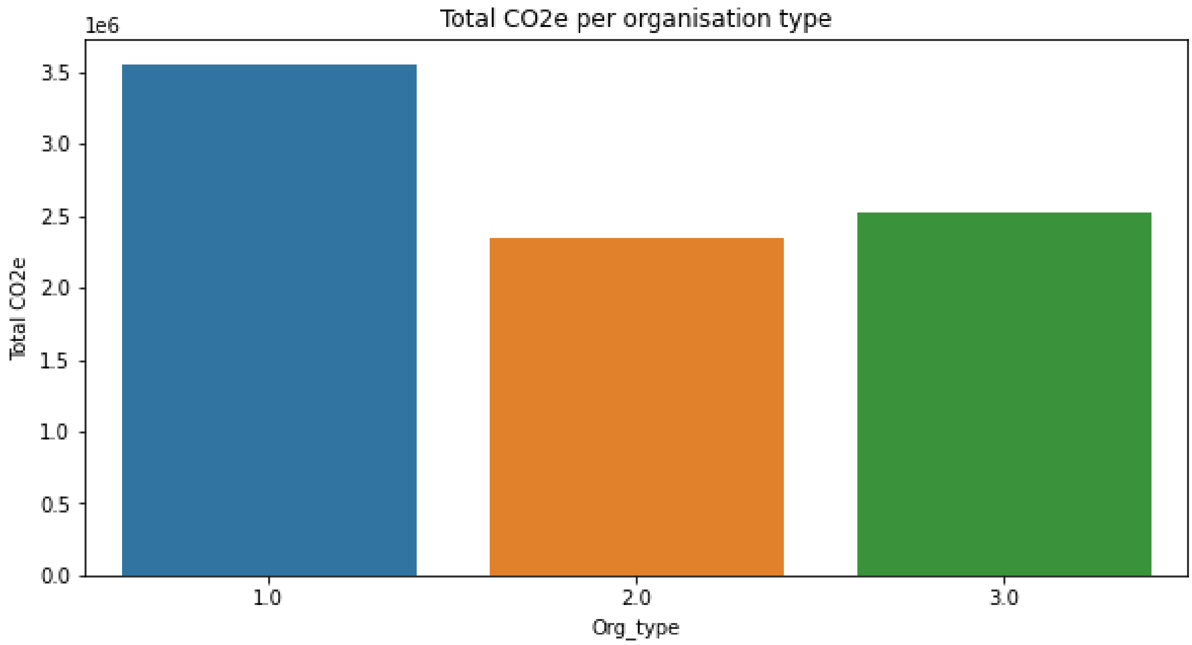

Figure 11:The plot of Total carbon emissions(kgCO2e) in the year 2018-19 by each organisation type- theatre companies (blue), art centers(orange) and touring organisations(green)

This is due to the obvious factors contributing to the emissions such as the amount of office area used, the number of staff members in office, the daily power, gas and other utility consumption of them as they are not responsible for emissions associated with running venues. These factors all follow the same trends as the amount of emissions as show in the plots for the same grouped by type of organisation (Figures 12-14). Any changes brought about in these factors can also bring about a positive/negative change in the response variable (Total carbon emissions). Hence it can be concluded that all of them (employees, organisation type, gas carbon emissions, floor area, electricity emissions) are linearly related to each other with some factor that can be found from the linear model later.

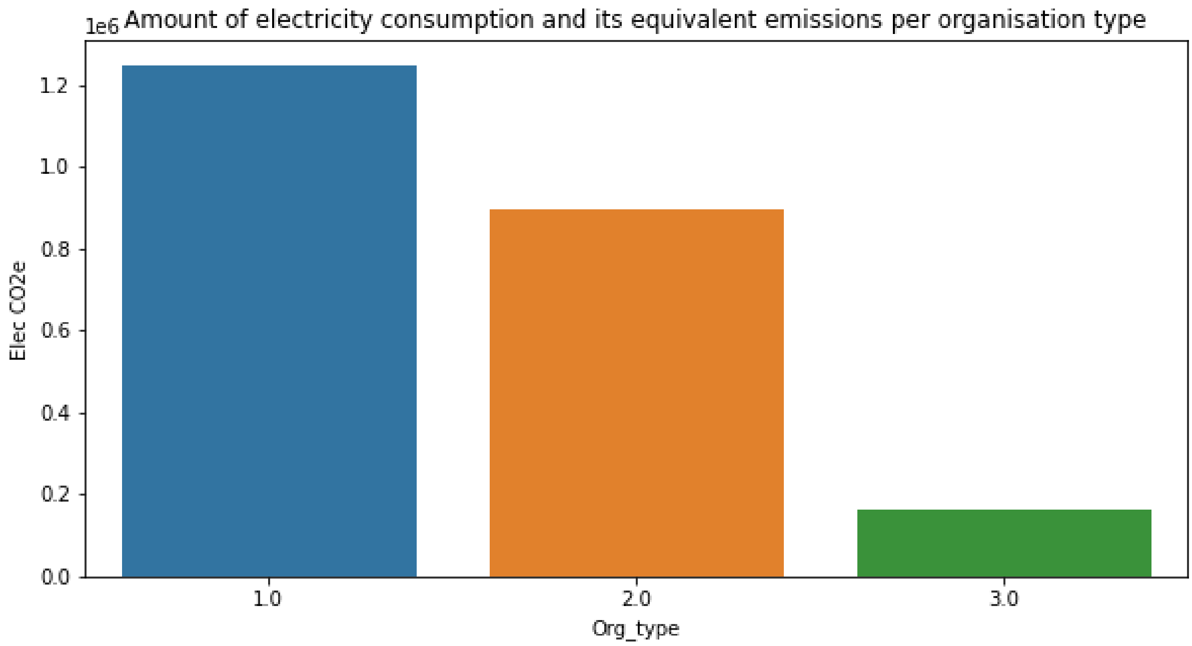

Figure 12: The total electricity consumption and the equivalent carbon emission in kgC02e in the year 2018-19 by each organisation type- theatre companies (blue), art centers (orange) and touring organisations(green)

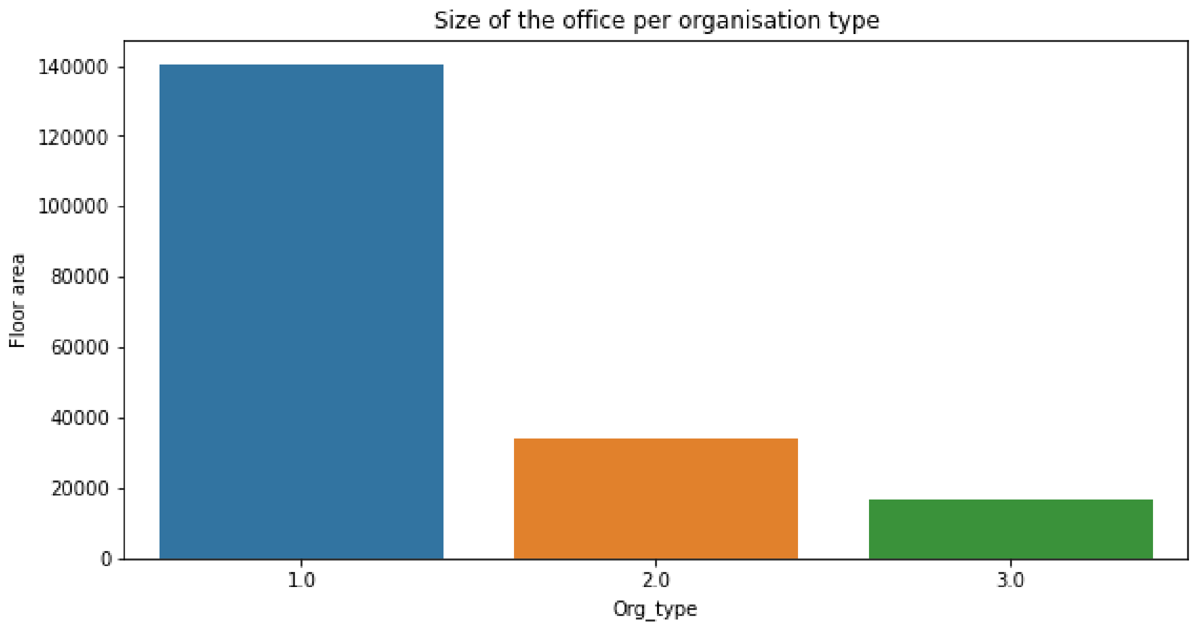

Figure 13:The plot of office size in area working with each organisation type- theatre companies (blue), art centers (orange) and touring organisations(green) in the year 2018-19

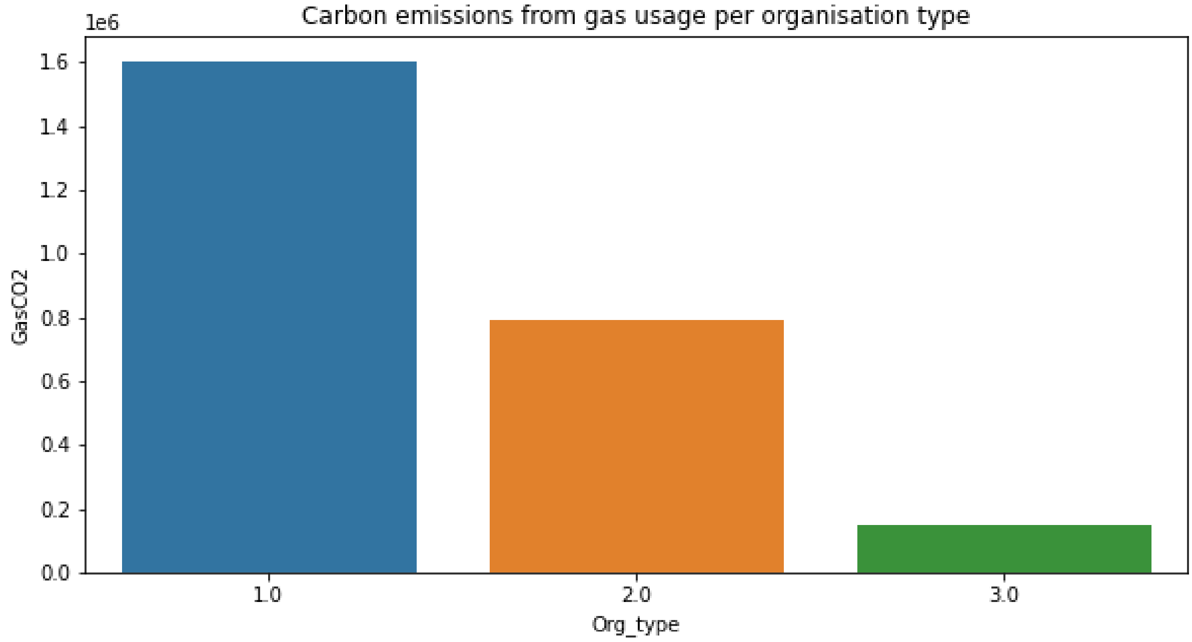

Figure 14: The plot of gas utilization carbon emissions in each organisation type- theatre companies (blue), art centers (orange) and touring organisations(green) in the year 2018-19

As shown below in Figure 15, the change in the number of employees has a different effect on the carbon emissions of a specific organisation for different types of organisations. It can be implied that, even with the increase in employees for a touring organisation the carbon emissions do not rise by a lot. This is majorly because the carbon emission for these type of companies is dominated by the travel emissions as they are not responsible for emissions associated with running venues.

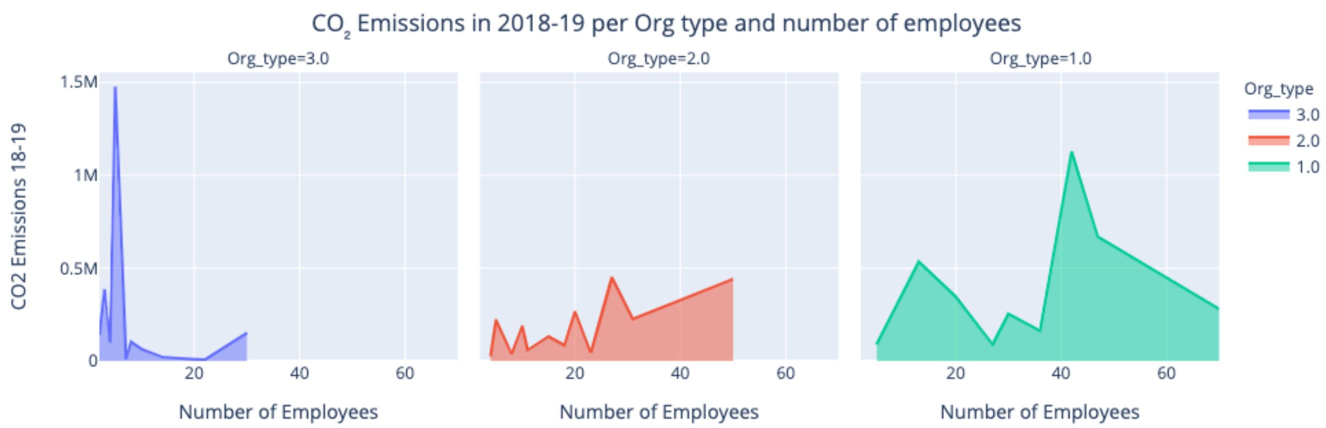

Figure 15: Relation of change in number of employees of an organisation(Green = theatres, Orange = art centers, Purple =touring org) and change in the total carbon emissions(kgCO2e)

Next, we can see the trend for electricity emissions that depends on the amount of electricity consumed by the people and the size of the office that runs the firm. Thus, we can see that the highest effect is on the theatre companies that run the entire show. The stage lighting, setup, equipment and the employees use up the most electricity at the venue. The larger the venue, the greater the audience accommodated, and thus the higher the electricity consumption. As concluded from results in the two figures, Figure 16 and 17 , with increase in floor size as well as number of employees for a theatre and art centre, the emissions related to electricity consumption also increase proportionately. From figure 16 we can observe that for tenant organisations the electricity usage is less as compared to the others because of small or nearly zero offices to run the firm. For art centres it is also because of the number of organisations and employees that the consumption is less but it still increases with increase in floor area.

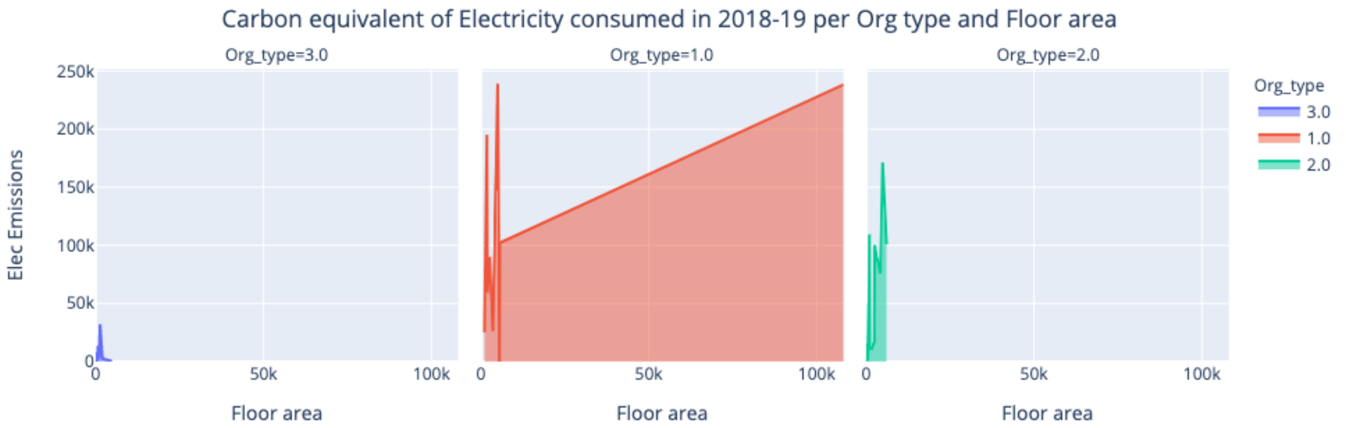

Figure 16:Change in floor area of an organisation(Green = art centers, Orange = theatres, Purple =touring org) affecting the electricity carbon equivalents(kgCO2e)

Humans are the biggest reason for consumption of resources (both renewable and non-renewable). The larger the population of a country the more is the required resources. This is true in all cases. As shown in Figure 17, with the increase in the size of offices for theatres (type1) and art centres (type 2) there is a proportionate increase in the electricity usage and emissions from the same. For the tenant groups there is a change with number of staff too but the change is very less due to other factors such as usage for set and equipment at venue.

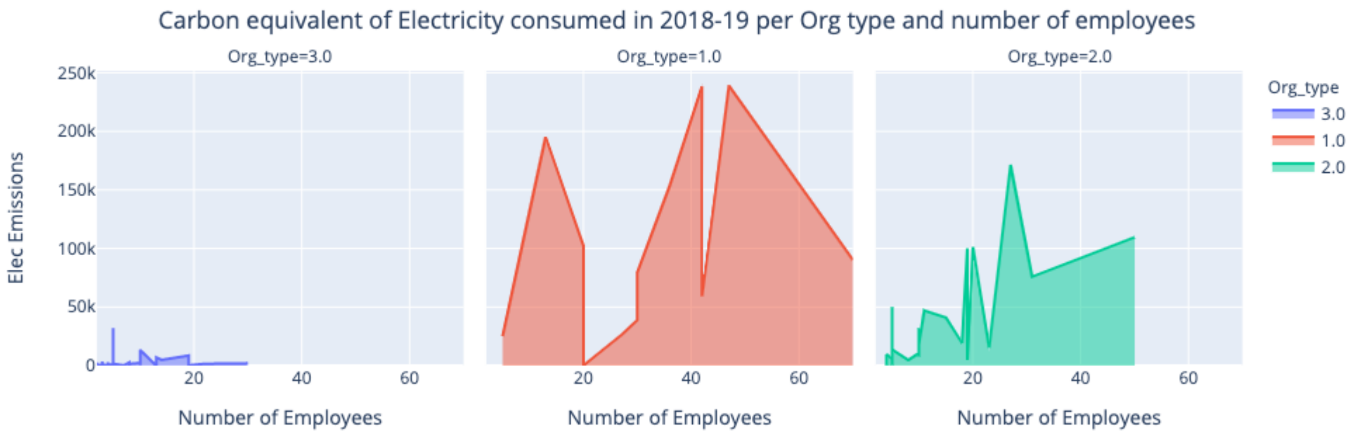

Figure 17:Change in number of employees of an organisation(Green = art centers, Orange = theatres, Purple =touring org) affecting the electricity carbon equivalents(kgCO2e)

We can also see the relation between electricity carbon equivalents and total carbon emissions of any organisation by plotting the scatterplot for the two variables Total CO2e and Elec CO2e from the given data (Figure 18) .

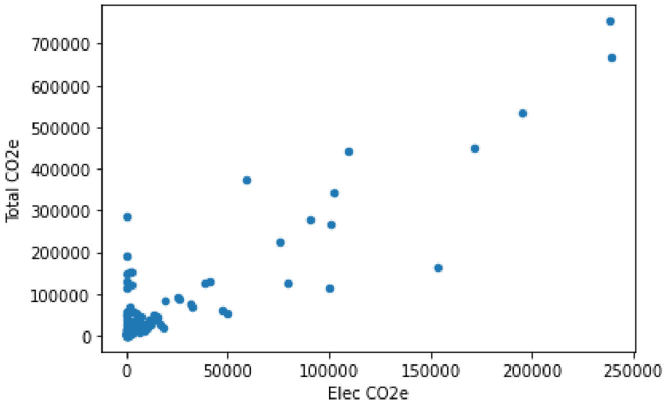

Figure 18: Linear Relation between the Total Carbon Emissions(y-axis) of any organisation and the total Electricity Carbon equivalents (x-axis)

Similarly, to show that total emissions have the same relationship with the floor area and number of staff members and show presenters as the electricity equivalents we can plot scatter plots for total emission against employee count and floor area, as well as plot the electricity component of emissions against the same parameters and compare them side by side to get better picture. Figure 19 and Figure 20 show that the floor area has a similar relationship with both the electricity emission from the usage and the total emissions from all components in the organisations run.

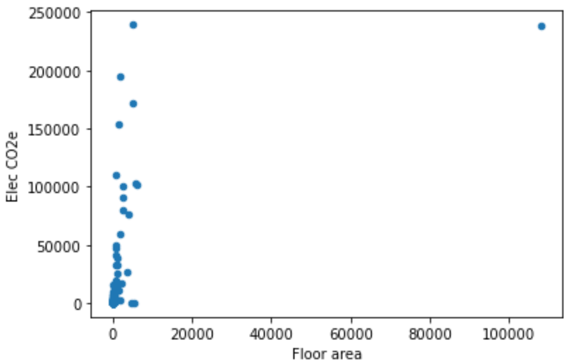

Figure 19:Electricity Carbon equivalent(y-axis) vs

The Floor area of the office (x-axis) of all organisations

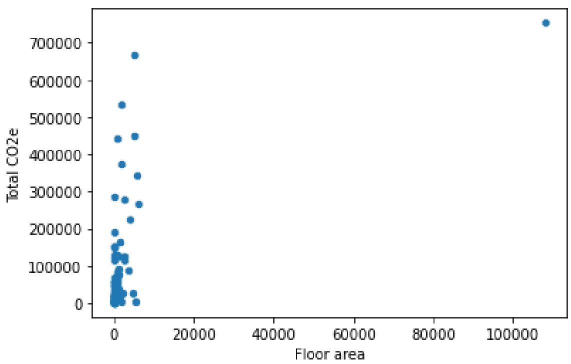

Figure 20: Total Carbon emission(y-axis) vs

The Floor area of the office (x-axis) of all orgs

Figure 22 and Figure 22 show that the human resources size also has a similarity in relationship with both the electricity emission from the usage and the total emissions from all the processes involved in driving an organisation.

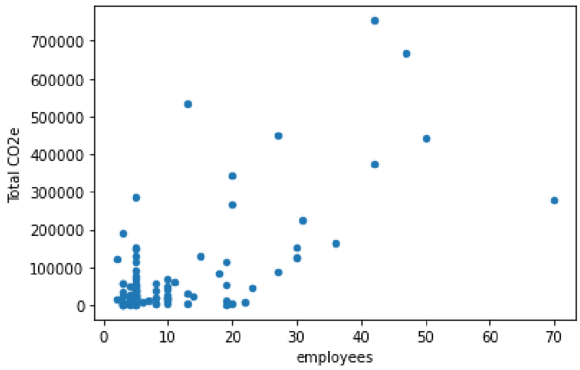

Figure 21:Electricity Carbon equivalent(y-axis) vs

The count of employees (x-axis) of all organisations

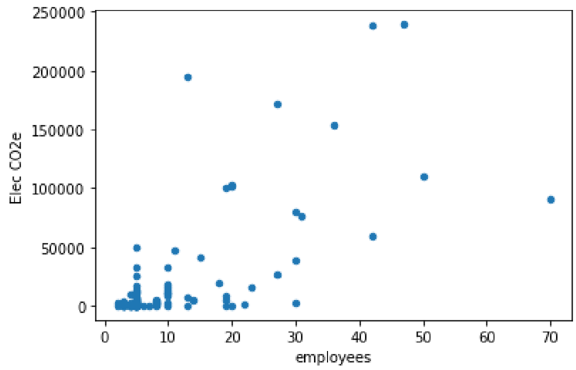

Figure 22: Total Carbon emission(y-axis) vs

the count of employees (x-axis) of all orgs

Regarding travel emissions, we have different cases where the domestic audiences and performers travels by plane, bus, train and private car and their respective CO_2 emissions are recorded as follows again grouped by the organisation type. It is evident from the visualisation on travel emissions of different transportation modes used by these organisations in Figures 23-27 that the touring/tenant organisations (type 3) emissions are highly dominated by travel because they have small or nearly no offices to run the production or other usage by staff and in-house team. However due to the dominant level travel in case of type3 organisations is high even with less number of employees for the touring organisations. The bus travel is very high for touring organisations the reason that can be assumed is that the need to move with a lot of equipment and performance set. Their target audience is also more of school students or universities which have bus as the major transportation compared to the older mature audiences. We can illustrate these conclusions further and help us visualise better.

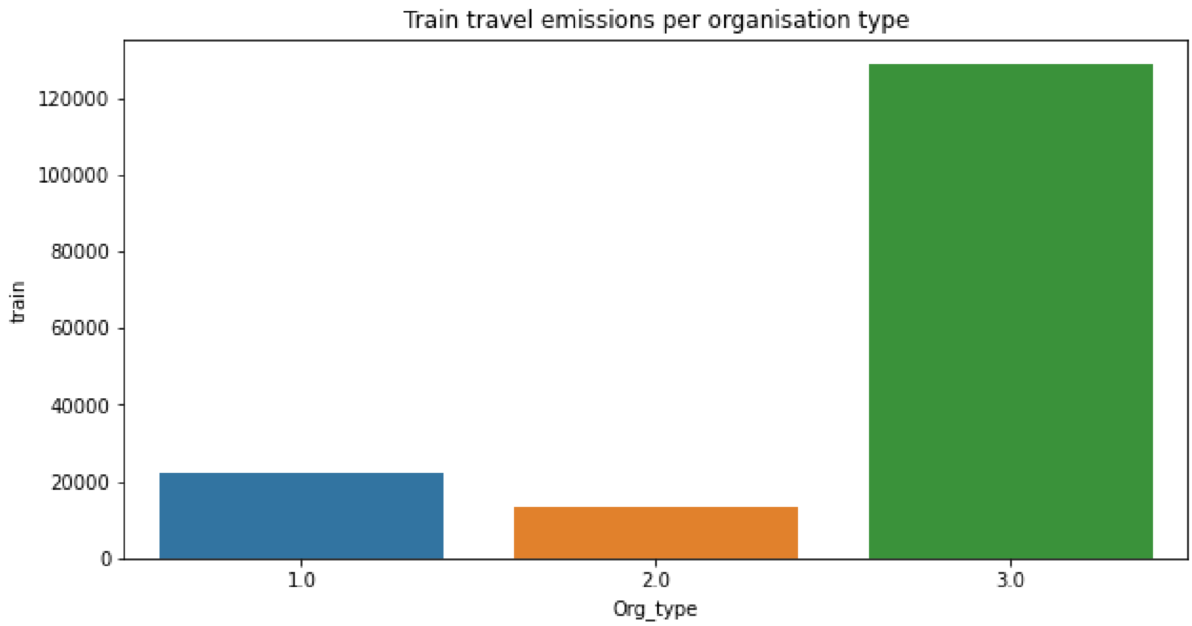

Figure 23: Amount of carbon emissions(kgCO2e) by Train travel (y-axis) according to organisation type (Blue= Theatres, Orange = Art Centers, Green = Tenants)

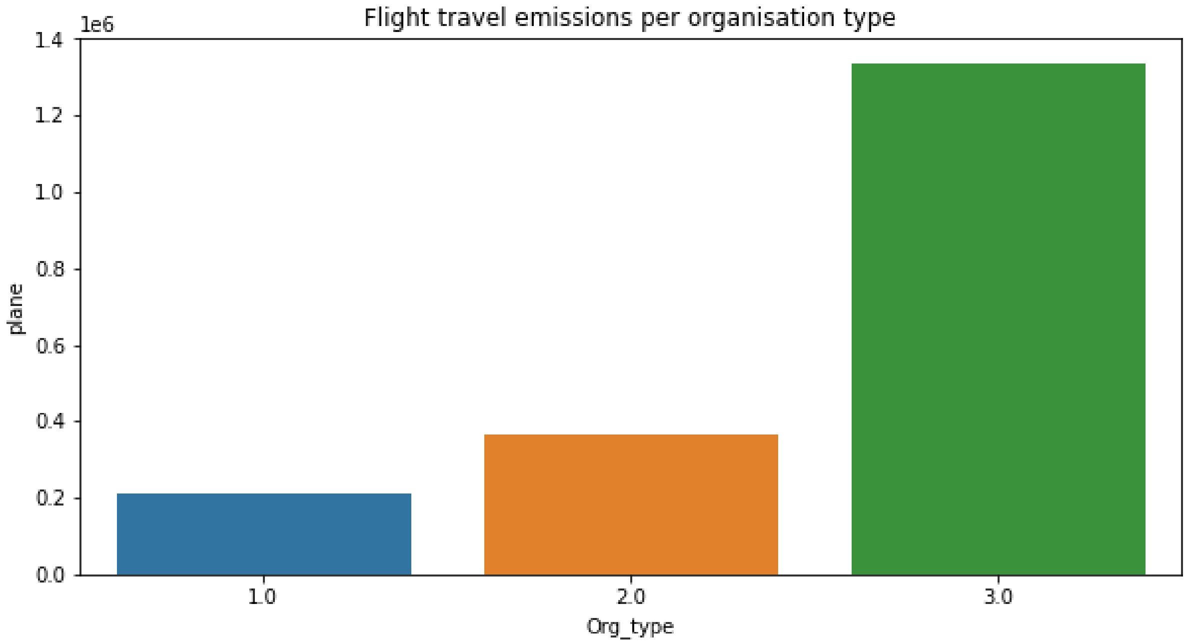

Figure 24:Amount of carbon emissions (kgCO2e) by Flight travel (y-axis) according to organisation type (Blue= Theatres, Orange = Art Centers, Green = Tenants)

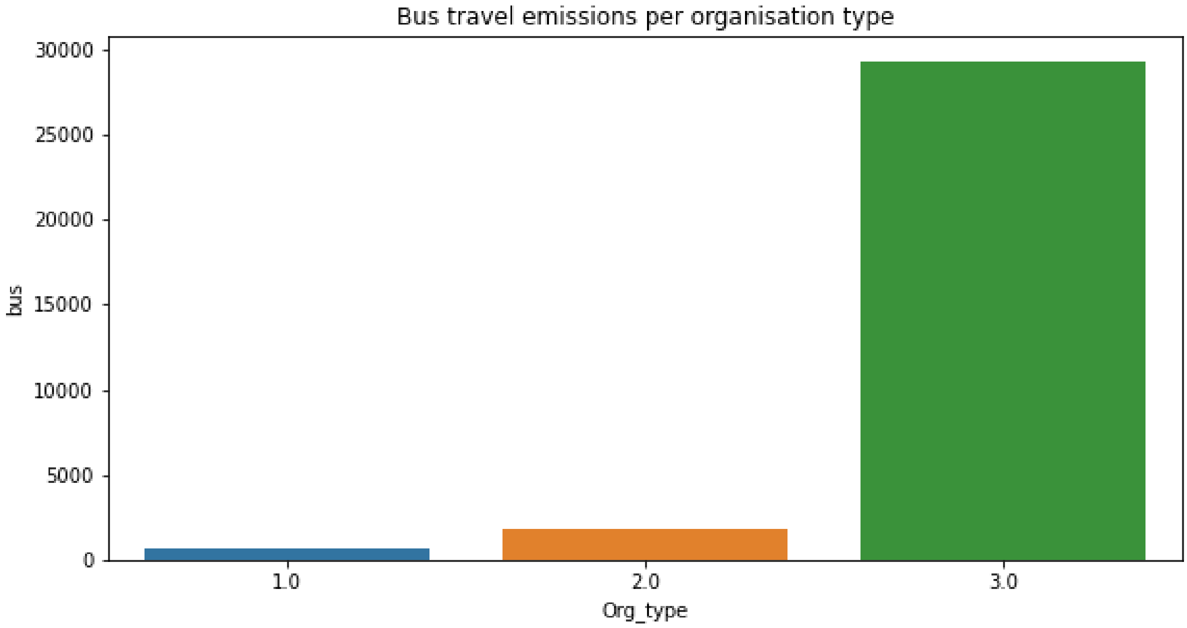

Figure 25: Amount of carbon emissions (kgCO2e) by Bus travel (y-axis) according to organisation type (Blue= Theatres, Orange = Art Centers, Green = Tenants)

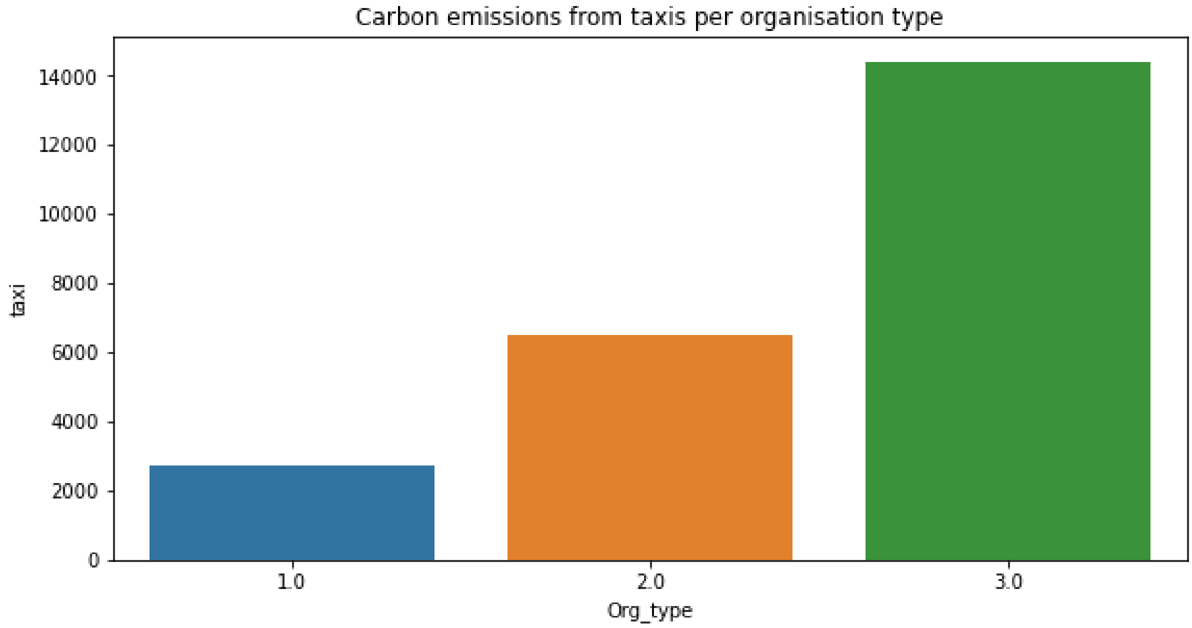

Figure 26:Amount of carbon emissions (kgCO2e) by Taxi travel (y-axis) according to organisation type (Blue= Theatres, Orange = Art Centers, Green = Tenants)

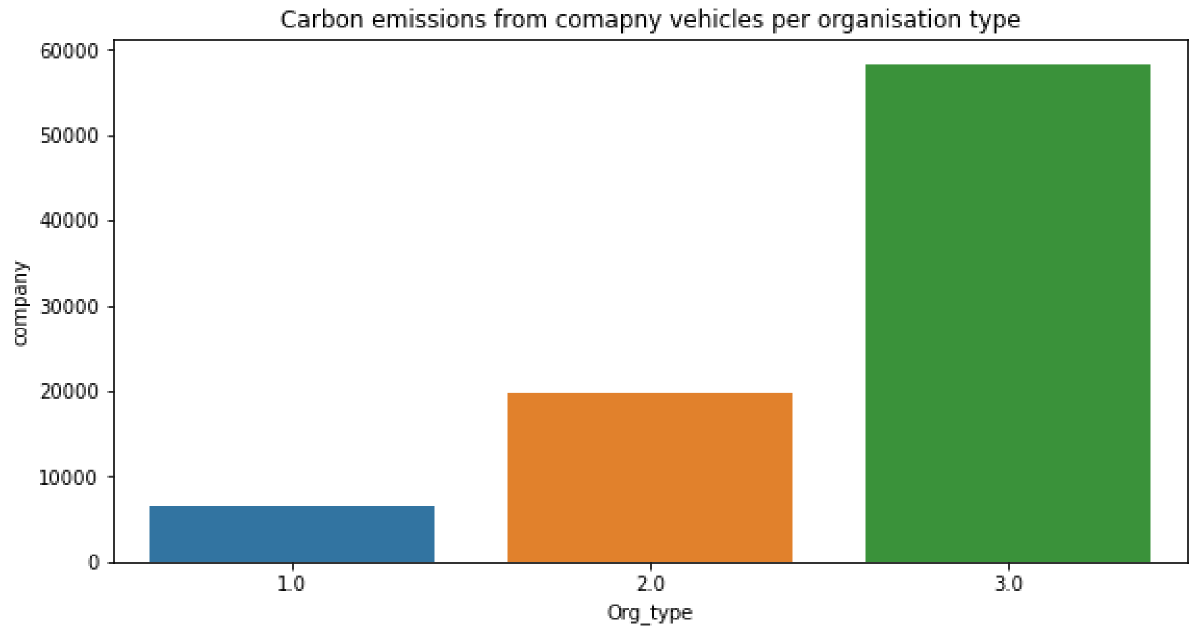

Figure 27: Amount of carbon emissions (kgCO2e) by company vehicle travel (y-axis) according to organisation type (Blue= Theatres, Orange = Art Centers, Green = Tenants)

Thus, we can say that if the travel components of emissions is completely removed in case of digital performances then only the electricity component plays the role of emitting the carbon. Also, the factors such as staff number, size of organisation and in turn the organisation type help us decide the electricity component from the given datasets. We now export this cleaned and processed dataset from Python to and excel sheet for further modelling in R.

Linear Regression – Electricity Consumption:

Using RStudio Version 1.4.1717, we can fit the model that has the total carbon emissions provided by the other components of carbon equivalents. First we look at the type of the variables. All variables, except the Org_type are quantitative. Thus, to factor the variable as qualitative data we use the command `factor()`. This will automatically tell R to use it as categorical variable while performing the linear regression. Our live performance model is given by the `lm()` command in R as follows:

```
live.model<- lm(`Total CO2e` ~ `Elec CO2e` + WaterCo2 + GasCO2 + LPGCO2 + FueloilCO2 +
                 `BIO CO2` + train + plane + bus + taxi + mileage + company +
                 `landf CO2` + `DMR CO2` + `Food CO2` ,data=live_df)
```

Figure 28: The above code creates a linear model named live.model which has Total CO2e as the response variable and all other variables like `Elec CO2e`, WaterCo2, GasCO2 , LPGCO2 , FueloilCO2 , ' BIO CO2` , train , plane , bus, taxi ,mileage, company , `landf CO2`, `DMR CO2` ,`Food CO2` from our dataset live_df which was extracted from python.

This is essentially a sum of all the given components including carbon equivalents of electricity, gas, water, travel, food waste, recycling and landfills used. Thus, this model will have a R-squared value of 1, which is ideal. We can try the components of actual consumption of theses utilities but it would not make much of a difference since all the carbon equivalent components are calculated using the carbon emission factors on these values, so it will just be a multiplicative factor from the live performace model (Appendix A).

In case of digital performances, electricity being the only important component, we can fit a model using R that essentially gives the estimates of the carbon equivalents of electricity consumption (Elec CO2) from the given dataset. We use the same entire model and use Elec CO2 as the y variable this time and the model obtained gives us the summary as shown in Appendix A. The model is given by code lm1 (Appendix B).

But as discussed with a creator of Big Data Show at Civic Digits, if we move to digital streaming performances, travel will essentially be decreased to negligible since performers and audiences can be where they are. Thus, travel components of emission will also be removed and prove to be insignificant. As shown in the model summary from Appendix A we can see that the p-values given inn the last column for all these components is way above the desired level ($p < 0.001$) for a coeffiecient to be significant. Thus, from the model lm1 cannot be considered a good model for the Elec CO2e estimations. Thus, we can apply simple step-wise selection method to select variables or covariates that are most significant in the estimation of the electricity consumption or emissions for a performance. The step-wise AIC selection method will create a regression model from a set of candidate predictor variables by entering and removing predictors in a stepwise manner based on Akaike Information Criteria (AIC) until there are no variables left to enter or remove (CRAN, 2021). After using the stepwise selection we can find that the model obtained is given as df_select2 in the R code from Appendix B.

Since, Org_type being a categorical variable, we factor it using the factor() function that is used to enumerate or encode as a numeric variable, any categorical variable that is an impacting variable and needs to be considered in the linear modelling Looking at the summary details of this model shown in Appendix A, we can see that the intercept of this model is 26610 i.e. if all the variables that impact the resulting electricity emissions are considered to be removed from the calculation then the total electricity emissions would still be 26610 kg CO_2e. Which can be influenced by some other factors that is not included by the function. It is sometimes very irrelavant to find an intercept, since if there is no organisation and no employees the production will not happen and thus, no point of estimating the emissions. However, the intercept is important to calculate the predicted values especially in the industry like analytics and market research and it is advised not to cross it out completely from the analysis. The p-values of all covariates given in the last column is significant to the dependent variable- Elec CO2e, thus we can say that the Org_type, employees, Floor area and GasCO2 are the variables that are significant predictors of the Elec CO2 variable. The estimates columns shows nothing but the coefficients of each independent variables selected in the model and their relationship can be explained in an equation as-

$$\begin{aligned} Elec\ CO2e(y) = {} & 26610 + 0.4794 \times (GasCO2) - 14700 \times \left(Org_{type2}\right) \\ & - 29020 \times \left(Org_{type3}\right) + 611.8 \times (employees) - 0.5282 \times (Floor\ area). \end{aligned} \tag{5}$$

An interpretation of the linear model for estimating the carbon equivalent of electricity is explained by equation (5) i.e., for every kgCO_2e increase in the gas emission carbon equivalent (GasCO2) the electricity equivalent of carbon emission(Elec CO2e) increases by 0.4794 kgCO_2e. Each employee in the organisation acounts to an increase in electricity emissions by 611.8 kgCO_2e. Similary, with every square meters of floor area change the electricity emissions have an impact by a fator of 0.5282 kgCO2e. Since, organisation type is a variable that is qualitative in nature, R picks the first group that is type 1 for theatres as the reference group. So, coeffiecient for Org_type2 in the above equation is the average difference of electricity emissions produced by theatres and art centres i.e. the theatres produce 14700 kgCO_2e more than art centers and coeffiecient of Org_type3 shows that tenant organisations produce 29020 kgCO_2e less electricty emissions that theatre companies. The R-squared value is not considered as an authentic method to compare model accuracy as it increases proportionately as the number of variables increases so we use something called the Adjusted R-squared. The adjusted R^2 of our digital performance model is 0.837, which means that 83% of the variation in electricity emissions can be explained using the significant factors which are the organisation type, size of office, staff count and the carbon emissions from the gas utilisation of the organisation which is considered fairly reliable model for these organisations.

The Linear Model Assumptions for Digital and Live Models:

From our model when we plot the Q-Q plot for model having Total CO2e as the dependent variable we can find that the residual errors in the model are not normally distributed and has Short tails normality with S-shape and points lying above the line in the left part of the plot and points below the line on right side of the plot, as shown in Figure 29.

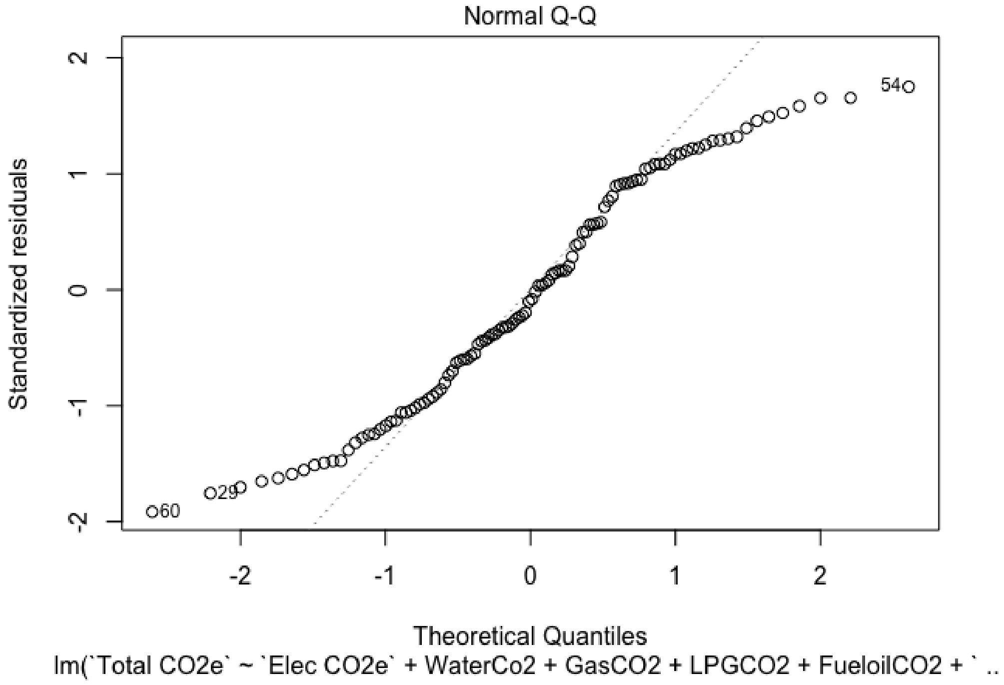

Figure 29: Normal Q-Q plot for the Live Performance Model

Now, plotting the Normal Q-Q plot for the digital performance model which has electricity as the dependent variable, we again see deviation from normality in data but this plot shows Heavy Tails normal behaviour and points lying below the line in the left part of the plot and points above the line on right side of the plot as given below (Figure 23).

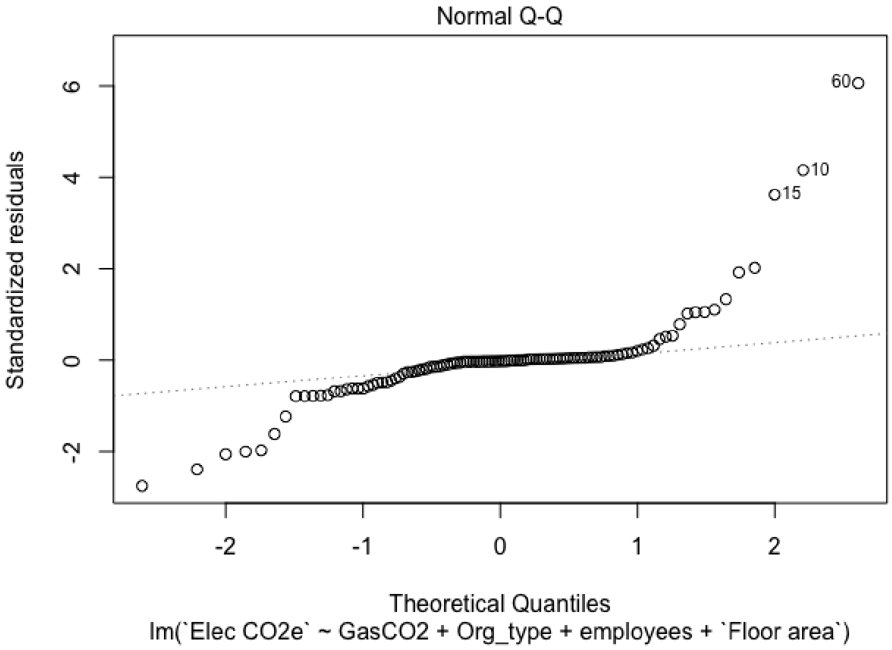

Figure 30:Normal Q-Q plot for the Digital Performance Model

For the live performance model, when we plot the residual vs fitted plot in R to check the variance of the errors and independence assumption, it shows the residual spread as shown in Figure 31.

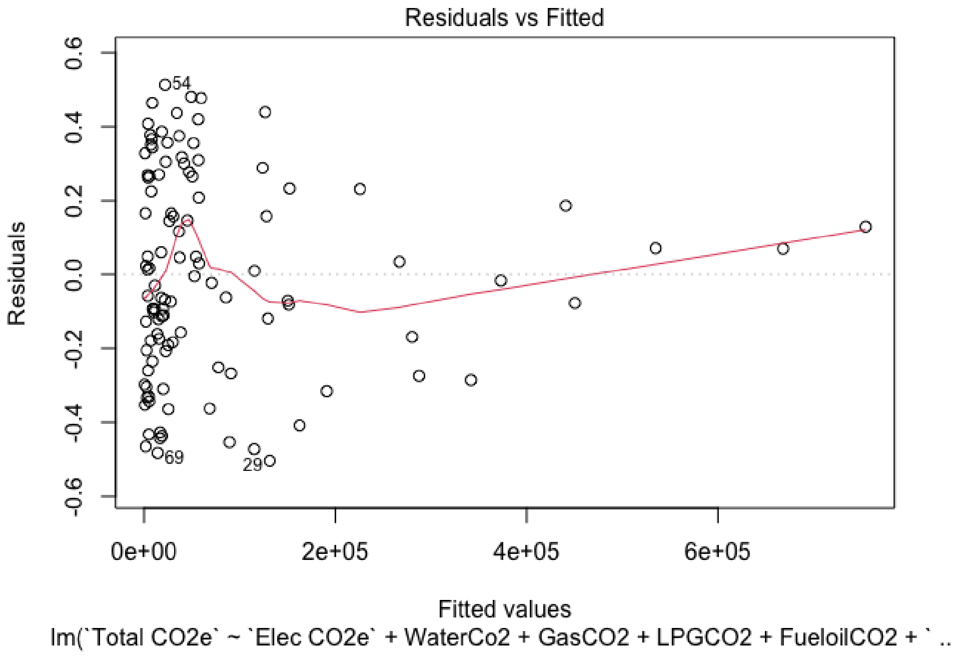

Figure 31: Independence Assumption check for Live Performance model

Similarly, for the digital performance model the spread is as shown in Figure 25.

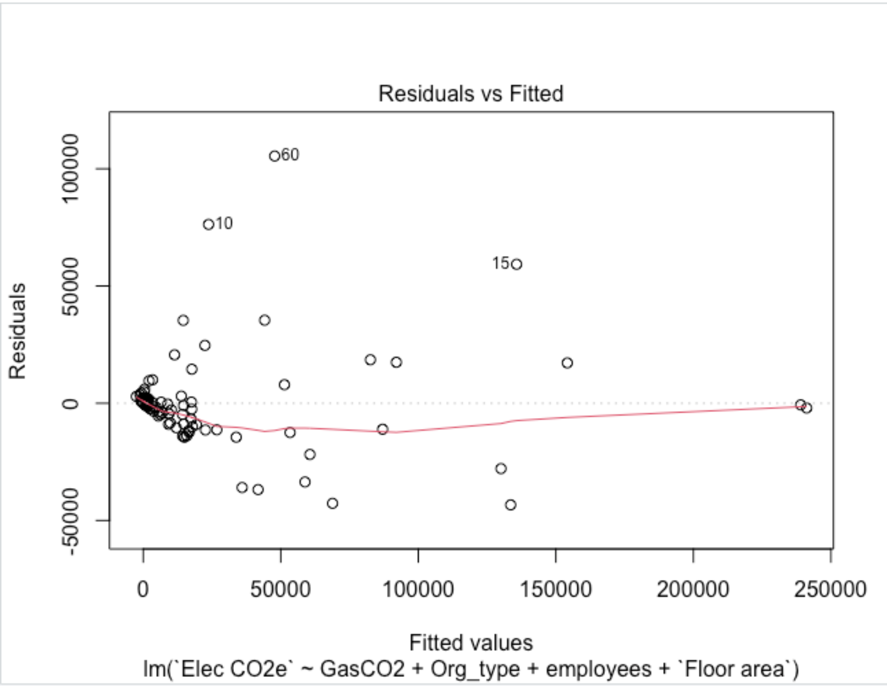

Figure 32: Independence Assumption check for Digital Model

Scatterplots should be created for each independent and dependent to determine if the relationship is linear (scatter forms a rough line). On scatterplots, binary variables can be distinguished by different markers, which aids in the investigation of patterns within groups.

Figure 33 shows the scatterplots for the variables that are considered for the estimation of final digital performance electricity emissions.

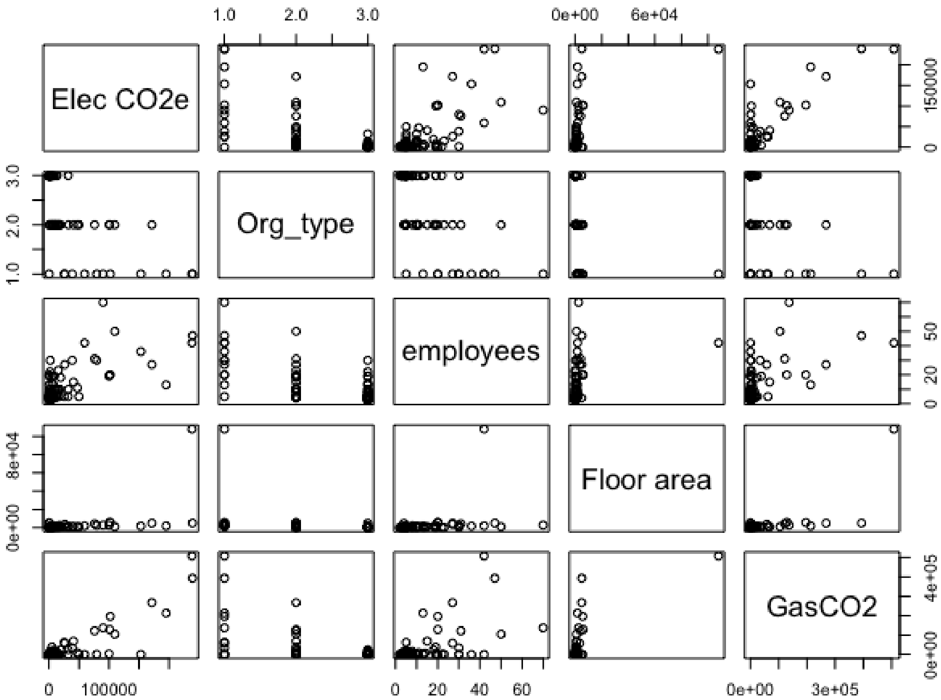

Figure 33: Scatter plots of all the components considered in the final Digital model to estimate the Electricity component of emissions

According to the Mayor of London's Office's Green Theatre Report, stage technology accounts for only 10% of emissions. However, front-of-house emissions accounted for 35% of total emissions, while heating and cooling rehearsal spaces accounted for 28% (Garrett, 2012). This also proves why gas utilities are an important factor too.

Let us compare our results from the calculations of for online performances by Civic Digits to the organisation that produces the highest amount of carbon emissions in the year 2018-19, in the entire running, performing and travel related to the show just to consider the highest coverage of audiences. From the data provided, Aberdeen Performing Arts has the highest level of total carbon equivalent over the year 2018-19 for in-person performance i.e., 754119 kgCO_2e and the electricity emissions equivalent of 238310 kgCO_2e annually. In their annual report we could find that the number of shows performed were 685 with over 320,000 audience members across 3 different venues. (Bremmer, 2019) The footprint estimated for a year of digital performances by Civic Digits for 2920 shows and approximately 2,920,000 audiences and over different venues all over Scotland from the above approach is giving us a highly reduced value of 82000 kgCO_2e per annum. This shows that by the digitisaion of performances the electricity emissions can be reduced to 34-35% of existing figures and with eliminitaion of travel by audience and actors, there can be a reduction in total emissions to as low as 10% of current figures. The calculated carbon impact of an hour of video streaming is heavily influenced by the geographical location of consumption and shows variability from country to country, because video streaming emissions are calculated using country-specific electrical grid emission factor. This illustrates the critical importance of governments continuing to drive the decarbonization of electricity grids. Thus, for the same show to run in the United States the emissions would be higher due to reduced or no use of renewable energy and thus, a higher emission factor for electricity (Stephens et al., 2021).

2.6 Discussion

2.6.1 Case Study: The Big Data Show.

As part of the show's operation, audience members are asked to download an application *SWIPE* that will include games a few days before and during the duration of the show. ("The Big Data Show – Civic Digits," 2021) The game is hosted on the show's website, which also streams the shows, and individual audiences can access and participate on their smartphones via the same app. The downloaded app can also control the phones of the audience members, and while it appears to be a simple addictive game, it is actually much larger and more complex. It tracks the audience's locations wherever they are using the app to play the game which is addictive. The website does not store any sensitive data anywhere on cloud, but it can instruct the user's phone to store and display data such as location data collected on the user's phone between the time the app is downloaded and the show begins. This location data is displayed on the phone screen as part of the show which is stored in encrypted form on the individual's device itself, but pretty much psyches out the player. In the theatre, the control system is used for both live and digital shows. In the online-only shows, the performances are uploaded through the laptops itself, so no external cameras or microphones are used. Nobody travelled for the performances, the performers as well as audiences were wherever they'd normally be. So travel emissions are reduced to zero. The shows are of 1hour and 10minutes duration with a plus-minus 5minute of delays due to gaming intermissions and entertainment between video segments of performances, which depends on the person running the show and the people playing the games. The shows can host audiences ranging from anywhere 500-1000 depending on complex interactive multiuser systems it is hard to predict what can be the limiting factor for a sum of maximum audiences, nor how far you can get when you tune them to work best under high loads. Thus, we can start with a show as soon as you finish the previous one which can theoretically be 20 shows in 24 hours and 7300 in the entire year. However, practically let us assume Civic Digits could do as good as 10 shows per day on weekends and 5 shows per day on weekdays, i.e. on an average 8 shows a day which brings the total to 2920 shows a year. This figure is much higher than the in-person shows that a company could do due to re-setup time and evacuating and re-entering audiences. Also, actors need breathers in between acts.

2.6.2 Online-only performances:

The carbon emissions for online shows completely eliminates the travel component. Due to pandemic and the work places shutting down and people working from home, the carbon emissions from biomass, recycling, food waste and landfill also is avoided. The only important factor that stays put in the calculations is the electricity consumption. To estimate the power consumption by Civic Digits for an online-only show we can consider all extra required equipment, data centres and transmission needed.

Let us consider a laptop-based video streaming which has minimal device overhead. The laptop is assumed to have internal microphone, speakers, camera and display, which are used in one of the video segment performance . However, as spoken by the co-creator of the Big Data Show, there is also one phone and one tablet generally connected to the sound system by Bluetooth, both used for incidental music and various sound effects during the show. One of them is controlled by an actor and other one by the technician running the show. The run time of the show is 70 minutes with participants (who need not travel anywhere), and 5 performers and 1 presenter who might or might not travel locally. We can consider it as an hour of streaming approximately.

The above-mentioned conventional approach from Carbon Trust is used to estimate the energy and emissions of video streaming per hour, and the analysis from the various literature is presented further. The breakdown of emissions and energy per hour of streaming from the results above yields the following using carbon emission factor of 0.26 $kgCO_2e$ /kWh ("Some-quick-carbon-conversions-2019-20," 2019) is given as follows. Data centres (including hosting, encoding and CDNs), account for less than 0.4 gCO_2e/hour and approximately 1.3Wh/hour, representing roughly 1% of total emissions and energy (Stephens et al., 2021). Network transmission (core and access) accounts for 6.2 gCO_2e /hour for average of high, standard and ultra 4K transmission on fixed network, 29.62 gCO_2e/hour. Home routers account for 18.7 gCO_2e /hour and 72Wh/ hour (Stephens et al., 2021). Finally, end-user devices, for a mix of devices used, account for 28 gCO_2e /hour (25 gCO_2e /hour from viewing devices and 3 gCO2e/hour from peripherals) and 96Wh/hour (86Wh/hour from viewing devices and 10Wh/ hour from peripherals) (Constable, 2011). The total emissions come something around 83 gCO_2e for one hour of video streaming per audience member. This account for one show by Civic Digits. However, for the considered average earlier, for 2920 shows and a maximum of 1000 audiences assumed per show, so the emissions from end-user devices would be 28 $kgCO_2e$ per show and 81,760 $kgCO_2e$ annually. While the other components will account for 160 $kgCO_2e$. Which gives us a total electricity equivalent of 82000 $kgCO_2e$. Also, comparing some results and conclusion from the Carbon Trust research to some daily-life activities we found that the total carbon equivalent emissions for microwaving a bag of popcorn for 4 minutes is 16 gCO_2e and the travel emissions from driving an average petrol car for 100 meters is 22 gCO_2e (Stephens et al., 2021). Thus, for the in-person performances, the audience travel will contribute to so much more than any amount of devices used for years of streaming.

There are many limitations because the data is not robust due to the source of data, which is a human survey from green champions of organisations. To make better calculations assumptions, we need a detailed analysis that includes the characteristics of each piece of equipment involved as well as the ability to consume power for each resource required. The conventional approach, on the other hand, is unsuitable for estimating the marginal carbon impact of a change in service level, such as a change in video quality (Stephens et al., 2021). In another discussion with an employee from Civic Digits, we found that it would be nearly impossible to determine the concise amount of power consumed by CODECs because each phone, tablet, or laptop that watches the stream will have its own codec, and codecs can be configured in so many different ways that you can't make any meaningful assumptions about power consumption unless you test a specific configuration. Even then, because they affect the bandwidth used and the amount of computation required by each end node, and they will all be following their own rules, it is a question that can only be answered in abstract terms. Also, the calculations and estimations made above are using a simple averaging assumption and multiplicative approach for all the devices and peripherals. With the exact values and configurations provided and also the availability of life-cycle assessment knowledge and data the estimations can further be accurate. Since, the data provided by Creative Carbon Scotland is based on a survey to the green champion of the companies, we cannot consider the reliability of the data. As a statistician it is important that the data used to create a model is reliable. Even though the accuracy of the model derived is high and considered reliable, we really cannot trust the data provided by man.

Ericsson's news podcast and their research work has cleared the myth of how using digital devices and various media does less harm to the planet than most of the other industries. They compared the electricity consumption of streaming a video and internet surfing to other general activities we conduct in our daily lives to get a better picture of the same. As per the results, power consumed in watching a streamed video for 2 hours on a TV screen with high end connections and data transmission rates is similar to that of running a new refrigerator for 24hours (Ericsson, 2020). Also, surfing the internet for 8 hours requires same amount of electricity as needed by an electric kettle to boil 1 litre of water (Ericsson, 2020).

In addition, they plan to present 'The Big Data Show' in a number of mid-size theatres to audiences of between 450-600 people. This show will last 90 minutes. The stage set is again minimal (dominated by a large screen) and show will involve a presenting team of 5 actors and technicians. The touring will involve a bit of travel factor added due to the team moving all over Scotland with no repeated venues, however less because it would involve just 2 vans and a small car. No use of extra cameras or microphones involved, but they use one radio mic for the performers. This can involve a little travel emissions but that is still very small compared to the amount of travel by audiences from around the world considering the audience reach of the online shows.

2.7 Conclusion

ICT industry is seen to have great potential in helping reduce the carbon emissions by as much as 15% (Ericsson, 2020). This can help decarbonize the world as all the sectors in the economy are digitalised. The Big Data Show by Civic Digits also requires the audiences download a series of videos as part of the game or quiz. The "Despacito" example, presented by Ericsson in their report, analyses the number of downloads and streams (over five billion) of the song 'Despacito' (released in 2017), and it consumed as much electricity as Chad, Guinea-Bissau, Somalia, Sierra Leone, and the Central African Republic combined in a single year. In the spring of 2018, this was stated in several media articles (Ericsson, 2020). More precisely, 5 billion downloads of this song to a smartphone require about 0.005TWh, when networks and data centres are factored for a single song download typically consumes 0.001kWh (Ericsson, 2020). Also, the results from their study concluded that the electricity needed for streaming a video for 2 hour on a laptop with high definition is almost equal to what is consumed by an electric car ride for a kilometre (Ericsson, 2020). The online digital footprint reduces the carbon footprint to such a high extent that in future, even if Civic Digits wants to conduct a tour around various educational institutes including school, colleges and universities around Scotland the carbon footprint would still be less compared to the number of audiences that can be reached. Due to streaming, the international travel can be avoided which reduces the major part of the travel emissions. The local travel involved with just the 5 actors and 1-2 technicians travelling in 2 vans and a car will still be negligible as compared to the hundreds and thousands of audiences travelling to the venues from around the world in different transportation modes.

The global transportation sector accounted for nearly a quarter of all anthropogenic carbon dioxide (CO_2) emissions in 2010, emitting 8.8 billion metric tonnes (Gt) of CO_2 and consuming 47 million barrels of oil per day (mbd) (Miller and Façanha, 2014). Thus, reduction or elimination of travel has helped reduced the carbon emissions to a large extent. Comparing the aviation industry emissions with the ICT emissions, it was also found that carbon emissions from a one person transatlantic return trip can account for emissions from 50 years of smartphone usage (Ericsson, 2020). Only 10% of the world's population uses aviation services (i.e. aeroplane travel) each year, with only the richest 1% being frequent flyers (Ericsson, 2020). This means that even if the sectors' footprints were of comparable magnitude, the impact per user would still differ significantly. However, because of the rapid nature of technological advancements and improvements in energy efficiency, it is difficult to project far into the future of the ICT sector. Post-pandemic and with growth of technology, emerging ICT trends add to the uncertainty of forecasting future emissions. Machine learning, blockchain and cryptocurrencies, IoT, and 5G mobile networks are among these trends. The impacts on the energy demand caused by these emerging trends is unpredictable and unclear as of now.

In conclusion, it is observed, derived and reviewed from literature as well that digitisation of the theatrical performance sector and art performances can lead to a high level of decarbonisation of the planet. The United Nations(UN) secretary-general Antonio Guterres has warned that the new report released by the UN body IPCC is a “code red for the humanity”, as it states that the environmental changes caused by human activities unequivocally are now “irreversible”, thus putting billions of people at risk (MassonDelmotte et al., 2021). Thus, switching to digital world, the theatre sector can contribute a lot towards eliminating carbon emissions from the unnecessary travel to venues, the setup, the equipment, the stage lights and generator fuel used and more. It might not be an ideal situation for the audience as they might not enjoy the feel of watching a performances comprising of a drama or a sketch scene on screens rather than witnessing live performance, but this involves doing a greater good to the world, so might as well be initiated before it's too late. As Greta Thunberg an environmental activist said in an interview with Reuters, "I hope that this can be a wakeup call, in every possible way" (Abnett, 2021).

From the results obtained from the regression model, we can conclude that factors that influence the electricity consumption of each of the organisations changes with the different types of delivery and art forms, the amount of floor area of each office, the staff and employees working for them and the utility of gas for heating and cooling and other important functions operated by them. The literature of digital performances was also reviewed and we can safely assume that the highest amount of carbon emissions of a streamed video performance comes from the electricity consumed by data transmission quality and rate, data storage and data centres, the user devices, the coding encoding devices and the video conferencing peripherals used. However, the amount of emissions produced by each of these depends highly on the specifications and configurations. For example, for high definition transmission and a TV screen used with emit much more carbon than a smartphone with standard definition. After comparing each of the results we have also drawn a conclusion that even though there is an increasing demand for technology and its substitutes, the total emissions with increased traffic in ICT will still leave enough room for the large amounts of emissions produced by in-person performances held by theatres, art centres, and various touring groups of these art presenters through the use of various energy and the amount of travel involved.

Electricity is not just the only component that impacts the footprint of an online streamed performance, but due to limitations on data availability and timeline restrictions we have limited the scope of our research to just electricity consumption from operational and embodied power of the user-devices, data centres and the equipment involved and data transmission, which play the most important role in a video streaming. This is sensitive to the characteristics of the transmission network also. These network characteristics include network equipment efficiency, network user population, and data traffic (Stephens et al., 2021). When comparing just the electricity component from the live performances in our regression model we could find that the type of the event company and in turn the office area, number of employees and the gas utilities make a significant amount of impact on the electricity consumption and thus, the carbon emission equivalent of the same. We analyse the difference in the amount of emissions due to electricity consumed by them and there is a vast difference shown in results. The reduction is as low as 10% of the current emissions from audiences travelling to the theatres and art centres to watch in-person shows. The reliability of data is also a point of question and the model is not considered very robust provided surveys are the source of data. It would have been a better and more reliable model if we had detailed information on the life cycle assessment of the various end-to-end data transmission, data storage, and streaming equipment, as well as details on the factors that influence power consumption, network characteristics such as transmission and quality of transmission, bitrate, definition used for streaming, and many more.

Appendices

- Appendix A

Live Performances Model:
Stepwise selection:

```
Step:  AIC=-253.69
`Total CO2e` ~ `Elec CO2e` + plane + GasCO2 + LPGCO2 + mileage +
    FueloilCO2 + train + WaterCo2 + `landf CO2` + company + bus +
    taxi + `BIO CO2` + `DMR CO2` + `Food CO2`

               Df  Sum of Sq        RSS     AIC
<none>                       8.0000e+00 -253.69
+ Org_type      1 0.0000e+00 8.0000e+00 -252.69
+ `Floor area`  1 0.0000e+00 8.0000e+00 -252.31
+ employees     1 0.0000e+00 8.0000e+00 -251.78
- `Food CO2`    1 3.5084e+05 3.5085e+05  917.44
- `DMR CO2`     1 3.1840e+06 3.1840e+06 1160.05
- `BIO CO2`     1 3.0037e+07 3.0037e+07 1406.92
- taxi          1 5.0554e+07 5.0554e+07 1464.19
- bus           1 6.8394e+07 6.8394e+07 1497.43
- company       1 7.9783e+08 7.9783e+08 1767.66
- `landf CO2`   1 2.4818e+09 2.4818e+09 1892.49
- WaterCo2      1 3.5698e+09 3.5698e+09 1932.48
- FueloilCO2    1 4.1041e+09 4.1041e+09 1947.83
- train         1 4.2362e+09 4.2362e+09 1951.31
- mileage       1 1.8346e+10 1.8346e+10 2112.54
- `Elec CO2e`   1 4.5987e+10 4.5987e+10 2213.63
- LPGCO2        1 8.5230e+10 8.5230e+10 2281.50
- plane         1 9.5146e+10 9.5146e+10 2293.60
- GasCO2        1 1.2475e+11 1.2475e+11 2323.40
Warning messages:
1: attempting model selection on an essentially perfect fit is nonsense
2: attempting model selection on an essentially perfect fit is nonsense
```

```
> summary(live.model)

Call:
lm(formula = `Total CO2e` ~ `Elec CO2e` + WaterCo2 + GasCO2 +
    LPGCO2 + FueloilCO2 + `BIO CO2` + train + plane + bus + taxi +
    mileage + company + `landf CO2` + `DMR CO2` + `Food CO2`,
    data = live_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.50385 -0.20183 -0.01989  0.25449  0.51299

Coefficients:
              Estimate Std. Error    t value Pr(>|t|)
(Intercept) -2.555e-02  3.821e-02 -6.690e-01    0.505
`Elec CO2e`  1.000e+00  1.377e-06  7.264e+05   <2e-16 ***
WaterCo2     1.000e+00  4.941e-06  2.024e+05   <2e-16 ***
GasCO2       1.000e+00  8.359e-07  1.196e+06   <2e-16 ***
LPGCO2       1.000e+00  1.011e-06  9.889e+05   <2e-16 ***
FueloilCO2   1.000e+00  4.608e-06  2.170e+05   <2e-16 ***
`BIO CO2`    1.000e+00  5.387e-05  1.856e+04   <2e-16 ***
train        1.000e+00  4.536e-06  2.205e+05   <2e-16 ***
plane        1.000e+00  9.571e-07  1.045e+06   <2e-16 ***
bus          1.000e+00  3.570e-05  2.801e+04   <2e-16 ***
taxi         1.000e+00  4.152e-05  2.408e+04   <2e-16 ***
mileage      1.000e+00  2.180e-06  4.588e+05   <2e-16 ***
company      1.000e+00  1.045e-05  9.567e+04   <2e-16 ***
`landf CO2`  1.000e+00  5.926e-06  1.687e+05   <2e-16 ***
`DMR CO2`    1.000e+00  1.654e-04  6.044e+03   <2e-16 ***
`Food CO2`   9.994e-01  4.981e-04  2.006e+03   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2952 on 94 degrees of freedom
Multiple R-squared:      1,     Adjusted R-squared:      1
F-statistic: 1.475e+12 on 15 and 94 DF,  p-value: < 2.2e-16
```

Digital Performances Model:
Estimate Elec CO2e, Including all variables:
lm1:

```
> lm1<- lm(`Elec CO2e`~ . ,data=digital_df)
> summary (lm1)

Call:
lm(formula = `Elec CO2e` ~ ., data = digital_df)

Residuals:
   Min     1Q Median     3Q    Max 
-43102  -5426    -68   1739 110109 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.458e+04  1.007e+04   2.441  0.01659 *  
Org_type2    -1.251e+04  8.918e+03  -1.403  0.16413    
Org_type3    -2.690e+04  9.491e+03  -2.834  0.00566 ** 
employees     5.142e+02  2.973e+02   1.729  0.08713 .  
`Floor area` -3.445e-01  3.150e-01  -1.094  0.27690    
WaterCo2      3.577e-02  3.331e-01   0.107  0.91471    
GasCO2        4.469e-01  5.363e-02   8.334 7.78e-13 ***
LPGCO2        3.573e-02  7.338e-02   0.487  0.62747    
FueloilCO2   -2.327e-01  3.099e-01  -0.751  0.45467    
`BIO CO2`    -1.313e+00  3.767e+00  -0.349  0.72824    
train         1.139e-02  3.047e-01   0.037  0.97025    
plane         7.362e-02  6.441e-02   1.143  0.25607    
bus          -1.523e+00  2.403e+00  -0.634  0.52784    
taxi          1.268e-02  2.782e+00   0.005  0.99637    
mileage      -3.437e-02  1.466e-01  -0.235  0.81511    
company       4.267e-02  7.054e-01   0.060  0.95190    
`landf CO2`   3.477e-01  4.447e-01   0.782  0.43641    
`DMR CO2`     1.928e+00  1.267e+01   0.152  0.87939    
`Food CO2`    1.980e+01  3.598e+01   0.550  0.58356    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 19760 on 91 degrees of freedom
Multiple R-squared:  0.85,	Adjusted R-squared:  0.8204 
F-statistic: 28.65 on 18 and 91 DF,  p-value: < 2.2e-16
```

Stepwise Selection:

```
Step:  AIC=2169.31
`Elec CO2e` ~ Org_type + employees + `Floor area` + GasCO2

                Df  Sum of Sq        RSS    AIC
<none>                        3.6867e+10 2169.3
+ plane          1 3.7304e+08 3.6494e+10 2170.2
+ `Food CO2`     1 2.1150e+08 3.6656e+10 2170.7
+ `landf CO2`    1 1.9556e+08 3.6672e+10 2170.7
+ FueloilCO2     1 1.3999e+08 3.6727e+10 2170.9
+ LPGCO2         1 6.2833e+07 3.6804e+10 2171.1
+ `BIO CO2`      1 5.3484e+07 3.6814e+10 2171.2
+ `DMR CO2`      1 5.1805e+07 3.6815e+10 2171.2
+ train          1 3.4382e+07 3.6833e+10 2171.2
+ taxi           1 2.8990e+07 3.6838e+10 2171.2
+ mileage        1 8.4746e+06 3.6859e+10 2171.3
+ bus            1 2.0019e+05 3.6867e+10 2171.3
+ company        1 8.0491e+04 3.6867e+10 2171.3
+ WaterCo2       1 1.4289e+04 3.6867e+10 2171.3
- `Floor area`   1 1.6260e+09 3.8493e+10 2172.1
- employees      1 2.6797e+09 3.9547e+10 2175.0
- Org_type       1 5.8328e+09 4.2700e+10 2183.5
- GasCO2         1 4.8663e+10 8.5530e+10 2259.9
```

Estimate Elec CO_2e, Including only selected variables from Stepwise selection: model.df:

```
> model.df<- lm(`Elec CO2e` ~  GasCO2 + Org_type + employees + `Floor area`
+                 ,data=digital_df)
> summary(model.df)

Call:
lm(formula = `Elec CO2e` ~ GasCO2 + Org_type + employees + `Floor area`,
    data = digital_df)

Residuals:
   Min     1Q Median     3Q    Max 
-43345  -4703   -421   1222 105441 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.661e+04  8.873e+03   3.000 0.003384 ** 
GasCO2        4.794e-01  4.105e-02  11.681  < 2e-16 ***
Org_type2    -1.470e+04  7.893e+03  -1.863 0.065282 .  
Org_type3    -2.902e+04  8.299e+03  -3.497 0.000693 ***
employees     6.118e+02  2.281e+02   2.682 0.008508 ** 
`Floor area` -5.282e-01  2.468e-01  -2.140 0.034669 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18830 on 104 degrees of freedom
Multiple R-squared:  0.8444,	Adjusted R-squared:  0.837 
F-statistic: 112.9 on 5 and 104 DF,  p-value: < 2.2e-16
```

```
> model.df<- lm(`Elec CO2e` ~  GasCO2 + Org_type + employees + `Floor area`
+                ,data=digital_df)
> summary(model.df)

Call:
lm(formula = `Elec CO2e` ~ GasCO2 + Org_type + employees + `Floor area`,
    data = digital_df)

Residuals:
   Min     1Q Median     3Q    Max
-43345  -4703   -421   1222 105441

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.661e+04  8.873e+03   3.000 0.003384 **
GasCO2        4.794e-01  4.105e-02  11.681  < 2e-16 ***
Org_type2    -1.470e+04  7.893e+03  -1.863 0.065282 .
Org_type3    -2.902e+04  8.299e+03  -3.497 0.000693 ***
employees     6.118e+02  2.281e+02   2.682 0.008508 **
`Floor area` -5.282e-01  2.468e-01  -2.140 0.034669 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18830 on 104 degrees of freedom
Multiple R-squared:  0.8444,	Adjusted R-squared:  0.837
F-statistic: 112.9 on 5 and 104 DF,  p-value: < 2.2e-16
```

- Appendix B:

RStudio Code:

```
#Dissertation
#linear model
#------Install all required Packages
library(ggplot2)
install.packages("MASS")
library(MASS)
install.packages("readxl")
install.packages("tidyverse")
library(readxl)
#Read dataset

df=read_excel("Carbon18-19.xlsx")
head(df)
live_df<- df[,2:20]
head(live_df)
digital_df<- df[,2:19]
head(digital_df)
digital_df$Org_type <- as.factor(digital_df$Org_type)
digital_df$Org_type
#selecting variables_ live
live_0 <- lm(`Total CO2e` ~1, data= live_df)
summary(live_0)
full.model_live<- lm(`Total CO2e`~.,data= live_df)
summary(full.model_live)
```

```
live_select<-step(live_0, scope=~Org_type + employees + `Floor area`
+ `Elec CO2e`
                  + WaterCo2 + GasCO2 + LPGCO2 + FueloilCO2 + `BIO
CO2` + train + plane + bus
                  + taxi +mileage+ company + `landf CO2` + `DMR CO2`
+ `Food CO2`, direction = "both")
summary(live_select)

live.model<- lm(`Total CO2e` ~ `Elec CO2e` + WaterCo2 + GasCO2 +
LPGCO2 + FueloilCO2 +
                  `BIO CO2` + train + plane + bus + taxi + mileage +
company +
                  `landf CO2` + `DMR CO2` + `Food CO2`
,data=live_df)
summary(live.model)
plot(live.model, which=2)


plot(live.model, which=1)

#selecting variables_ digital

full.model_digital <- lm(`Elec CO2e`~., data = digital_df)
summary(full.model_digital)
df_0.digital<- lm(`Elec CO2e` ~1, data= digital_df)

df_select1<-step(full.model_digital, scope=~Org_type + employees +
`Floor area` + WaterCo2 +
                GasCO2 + LPGCO2 + FueloilCO2 + `BIO CO2` + train +
plane + bus + taxi +mileage
                + company + `landf CO2` + `DMR CO2` + `Food CO2`,
direction = "both")
summary(df_select1)


df_select2<-step(df_0.digital, scope=~Org_type + employees + `Floor
area` + WaterCo2 +
                  GasCO2 + LPGCO2 + FueloilCO2 + `BIO CO2` + train +
plane + bus + taxi +mileage
                + company + `landf CO2` + `DMR CO2` + `Food CO2`,
direction = "both")
summary(df_select2)

#-- Linear regression model with only important variables

model.df<- lm(`Elec CO2e` ~  GasCO2 + Org_type + employees + `Floor
area`
              ,data=digital_df)
summary(model.df)

lm1<- lm(`Elec CO2e`~ . ,data=digital_df)
summary (lm1)
# Assumptions
#1. Residuals are normally distributed
plot(model.df, which=2)
```

```
qqnorm(resid(model.df))
qqline(resid(model.df))
#heavy tailed residuals

#2. Errors have constant variance

plot(model.df, which=1)
# plot(residuals(model.df)~predict(model.df))

# 3. Residuals are independent
acf(residuals(model.df))

# 4. true linear relationship
pairs(~`Elec CO2e` + Org_type + employees + `Floor area` + GasCO2,
data= digital_df)
```

Bibliography

Abnett, K., 2021. Thunberg: “Massive public pressure” needed to galvanize climate fight | Reuters [WWW Document]. URL https://www.reuters.com/business/environment/activist-greta-thunberg-now-plans-attend-un-climate-conference-scotland-2021-08-09/ (accessed 8.27.21).

About Creative Carbon Scotland | Creative Carbon Scotland [WWW Document], 2021. URL https://www.creativecarbonscotland.com/about/ (accessed 8.9.21).

About us – Civic Digits [WWW Document], 2021. URL https://civicdigits.com/about-us/ (accessed 8.9.21).

Andrae, A., Edler, T., 2015. On Global Electricity Usage of Communication Technology: Trends to 2030. Challenges 6, 117–157. https://doi.org/10.3390/challe6010117

Belenky, P., 2011. The Value of Travel Time Savings : Departmental Guidance for Conducting Economic Evaluations Revision 2.

Brander, M., 2012. Greenhouse Gases, CO2, CO2e, and Carbon: What Do All These Terms Mean?

Bremner, M., 2019. ABERDEEN CITY COUNCIL COMMITTEE Strategic Commissioning Aberdeen Performing Arts-ALEO Annual Report.

Butterworth, C., 2020. How using a CDN is better for people and the planet. [WWW Document]. WHOLEGRAIN digital. URL https://www.wholegraindigital.com/blog/how-using-a-cdn-is-better-for-people-and-the-planet/

Chouldechova, 2021. Regression diagnostic plots [WWW Document]. URL https://www.andrew.cmu.edu/user/achoulde/94842/homework/regression_diagnostics.html (accessed 8.23.21).

Cisco, 2019. Cisco Visual Networking Index: Forecast and Trends,2017-2022.

Constable, G., 2011. Power Consumption of Videoconferencing Equipment.

Costenaro, D., Duer, A., EnerNOC Utility Solutions, 2012. The Megawatts behind Your Megabytes: Going from Data-Center to Desktop.

CRAN, 2021. Variable Selection Methods [WWW Document]. CRAN. URL https://cran.r-project.org/web/packages/olsrr/vignettes/variable_selection.html (accessed 8.24.21).

Ericsson, 2020. A quick guide to your digital carbon footprint carbon footprint. https://doi.org/10.3390/su10093027

European Commission, 2021. Causes of climate change | Climate Action [WWW Document]. European Commission. URL https://ec.europa.eu/clima/change/causes_en (accessed 8.23.21).

Garrett, I., 2012. Theatrical Production’s Carbon Footprint.

Geist, J.-N., 2019. “Climate crisis: The Unsustainable Use of Online Video” : Our new report [WWW Document]. The Shift Project. URL https://theshiftproject.org/en/article/unsustainable-use-online-video/ (accessed 8.20.21).

Grce, S., 2021. The Carbon Footprint Of ... Video Streaming and Games? It’s A Thing | Nasdaq [WWW Document]. URL https://www.nasdaq.com/articles/the-carbon-footprint-of-...-video-streaming-and-games-its-a-thing-2021-01-19 (accessed 8.19.21).

IEA, 2020. Data Centres and Data Transmission Networks. Paris.

Kamiya, G., 2020a. The carbon footprint of streaming video: fact-checking the headlines – Analysis - IEA [WWW Document]. URL https://www.iea.org/commentaries/the-carbon-footprint-of-streaming-video-fact-checking-the-headlines (accessed 8.15.21).
Kamiya, G., 2020b. Factcheck: What is the carbon footprint of streaming video on Netflix? | Carbon Brief [WWW Document]. URL https://www.carbonbrief.org/factcheck-what-is-the-carbon-footprint-of-streaming-video-on-netflix (accessed 8.14.21).
Lenzen, M., 1999. Total requirements of energy and greenhouse gases for Australian transport. Transportation Research Part D-transport and Environment 4, 265–290.
Liu, C.-T., 2016. Correlation and Regression with R.
Malmodin, J., Lundén, D., 2018. The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015. Sustainability 2018, Vol. 10, Page 3027 10, 3027. https://doi.org/10.3390/SU10093027
Masso, G., 2021. Theatre company reveals digital output reduced carbon emissions by 98% [WWW Document]. URL https://www.thestage.co.uk/news/theatre-company-reveals-digital-output-reduced-carbon-emissions-by-98 (accessed 8.12.21).
MassonDelmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T.W.O.Y.R.Y. and B.Z., 2021. Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.
Miller, J.D., Façanha, C., 2014. THE STATE OF CLEAN TRANSPORT POLICY A 2014 SYNTHESIS OF VEHICLE AND FUEL POLICY DEVELOPMENTS.
Netflix, 2021. How to control how much data Netflix uses [WWW Document]. Netflix. URL https://help.netflix.com/en/node/87 (accessed 8.21.21).
Ong, D., Moors, T., Sivaraman, V., 2014. Comparison of the energy, carbon and time costs of videoconferencing and in-person meetings. Computer Communications 50, 86–94.
Ozawa-Meida, L., Brockway, P., Letten, K., Davies, J., Fleming, P., 2013. Measuring carbon performance in a UK University through a consumption-based carbon footprint: De Montfort University case study. Journal of Cleaner Production 56, 185–198. https://doi.org/10.1016/J.JCLEPRO.2011.09.028
Raghavan, B., Ma, J., 2011. The Energy and Emergy of the Internet.
Rawlings, J.O., 1998. Applied regression analysis : a research tool. Springer.
Ritchie, H., 2020. Climate change and flying: what share of global CO2 emissions come from aviation? - Our World in Data [WWW Document]. URL https://ourworldindata.org/co2-emissions-from-aviation (accessed 8.17.21).
Sausen, R., Schumann, U., 2000. Estimates of the Climate Response to Aircraft CO2 and NO x Emissions Scenarios. Climatic Change 2000 44:1 44, 27–58. https://doi.org/10.1023/A:1005579306109
Some-quick-carbon-conversions-2019-20, 2019.
Statista Reseach Department, 2020. Internetnutzer - Anzahl weltweit 2019 | Statista [WWW Document]. URL https://de.statista.com/statistik/daten/studie/805920/umfrage/anzahl-der-internetnutzer-weltweit/ (accessed 8.13.21).
Stephens, A., Tremlett-Williams, C., Fitzpatrick, L., Acerini, L., Anderson, M., Crabbendam, N., 2021. Carbon-impact-of-video-streaming.
The Big Data Show – Civic Digits [WWW Document], 2021. URL https://civicdigits.com/project/the-big-data-show/ (accessed 8.9.21).

The Paris Agreement | UNFCCC [WWW Document], 2021. URL https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (accessed 8.10.21).
Tripathi, A., 2019. Feature Selection Techniques in Regression Model – Data Science Duniya. Data Science Duniya.
Versteijlen, M., Perez Salgado, F., Janssen Groesbeek, M., Counotte, A., 2017. Pros and cons of online education as a measure to reduce carbon emissions in higher education in the Netherlands. Current Opinion in Environmental Sustainability 28, 80–89. https://doi.org/https://doi.org/10.1016/j.cosust.2017.09.004
Wiedmann, T., Barrett, J., 2011. A greenhouse gas footprint analysis of UK Central Government, 1990–2008. Environmental Science & Policy 14, 1041–1051. https://doi.org/10.1016/J.ENVSCI.2011.07.005