Economics of Animation: Role of the Animation Industry in Promoting Development

Presented in Partial Fulfillment of the Requirements for the Degree in Statistical & Data Sciences and Economics in the
MCS and Economics Departments at The College of Wooster

by
David Sokurov
The College of Wooster
2024

Advised by:

Moses Luri, Ph.D

THE COLLEGE OF

WOOSTER

Abstract

This research highlights the role of the animation industry in economic development, showcasing it as a potential route for new job opportunities and technological advancements. To specify this relationship, as animations become popular, various theme parks attract individuals by showcasing well-known and familiar characters. This attraction can extend beyond the border, drawing tourists from around the globe and thus stimulating the economy of the hosting country. Panel data spanning 18 years (2004-2022) across 7 countries and 19 theme parks was used to analyze this relationship, encompassing data from reputable sources such as The Worlds Bank, The Theme Park Database, and Google Trends. After several stages of analysis, this research concluded that animation popularity increases theme park visitation trends, which later attract tourists to various locations. Therefore, the animation industry could be a tool for stimulating economic development.

Acknowledgments

I want to acknowledge Professor Luri for his assistance and guidance throughout this project. I also want to acknowledge Professor Long for her instrumental role in expanding my understanding of economics. I attribute my acceptance to Georgetown University and the bestowal of an almost full-ride scholarship from Case Western Reserve University to her. Her guidance and mentorship have been vital in my personal and academic growth within the field, and I am deeply grateful for her support.

Contents

List of Figures

Chapter 1

Introduction

The animation industry has always been a window into another world, and I have always enjoyed watching and analyzing animations alongside various movies. I have always viewed this industry as a type of art that can tell a story and express information, as well as a wide range of emotions, in ways that all individuals can understand with ease despite their level of education. I find this field quite fascinating, and I have the utmost respect for the craft of storytelling and ingenuity that goes into it. The world of animations had a significant impact on my life as it taught me how to dream big and encouraged me to think outside the box, which contributed to why I ended up pursuing higher education halfway across the world today.

The history of this industry dates back to the early 20th century in France. The first animated film was released in 1908, called Fantasmagorie by Émile Cohl, and since then, the animation industry has grown exponentially, reaching the global market. One would dare to argue that animations have transcended many cultural boundaries and captivated audiences worldwide. Many animated movies such as Spirited Away, Grave of the Fireflies, Soul, and Toy Story, to name a few, have touched the hearts of many and gained numerous critical acclaims. On top of that, many animated movies tackle and touch upon many ongoing social issues and, with that, encourage discussions among many individuals. In a sense, one could say that animations serve as a tool to foster empathy and understanding among diverse communities. The industry has become a billion-dollar, multifaced industry, showing its ability to inform and educate its viewers.

This research analyzes the relationship between the animation industry's popularity and its influence on economic activity. It primarily focuses on the effects of film-induced tourism and how this phenomenon can enrich local businesses, create job opportunities, and stimulate the economy overall. By exploring this topic, I aim to bring to light how the popularity of different animations can serve as a catalyst for attracting tourists to a country. Through an empirical analysis, I quantify the contribution of the animation industry to increasing tourism rates and, consequently, to broadening economic activity.

The animation industry represents a dynamic and multifaced sector that encompasses the creation of moving images through a sequence of frames, giving viewers an illusion of motion. The concept of creating animated content involves techniques including traditional hand-drawn computer-generated images (CGI), stop-motion, and various other innovative methods. The products produced by this industry are used for many reasons, including different films, television shows, advertisements, video games, and many online platforms.

With the growth of digital platforms and streaming services, the animation industry has become a multi-billion-dollar industry that produces a wide range of content, including animated movies, TV series, video games, advertising campaigns, instructional materials, and virtual reality experiences. The market for this industry as of 2023 is equal to roughly 19.89 billion dollars, and on top of that, it is expected to increase by 2026 and reach an outstanding number of $27.10 billion US dollars (Statista). This goes hand-in-hand with the phrase by John Hawkins, which emphasizes that "Creative industries can create the value of USD 22 billion every day." Therefore, this industry has become quite a powerful component of the economy.

Coincidentally, film-induced tourism has become a novel phenomenon with the rise of digital platforms. It entails the idea of enthusiasts exploring various locations featured in their favorite movies: animated movies can inspire individuals to travel by showcasing captivating worlds and characters or enacting feelings of belonging and nostalgia. With that being said, the animation industry has contributed immensely to economic development through the growth of tourism rates by promoting cultural events such as theme park visitations, Comicon gatherings, and anime conventions. In other words, film-induced tourism has the potential to be used as a marketing tool to promote and sell various destinations by attracting people to sights associated with animated movies.

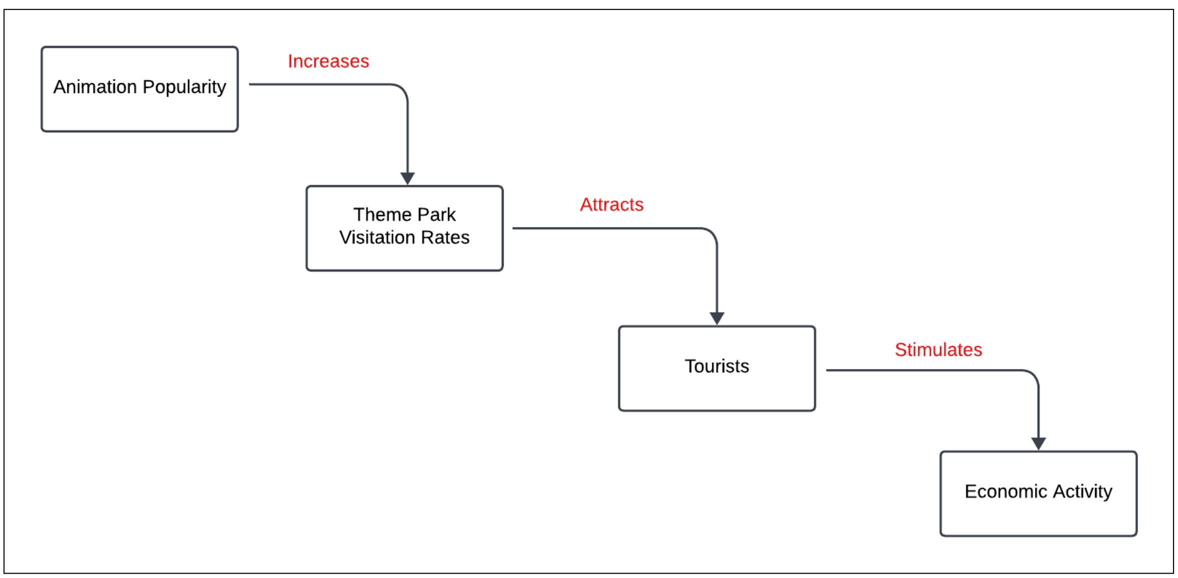


Figure 1.1: Economics of Animation Model: Unveiling Economic Growth Through Animation

Studying the impact of the animation industry on tourism rates is crucial for understanding the symbiotic relationship it has with economic development. As portrayed in Figure 1.1, this Industry can promote film-related tourism, which could impact target countries' economies, stimulate revenue, support local businesses, create employment opportunities, and infrastructural development. By analyzing in further detail how the animation industry stimulates economic development by boosting tourism rates, this research aims to determine whether the animation industry can foster economic growth and development in targeted countries or regions through animation-themed attractions from successful franchises. This research aims to prove whether the animation industry can be used as a tool to promote economic development in targeted countries.

In conducting this research, I used a panel dataset that captures both time and country effects and analyzed the Industry's impact on tourism rates across various countries over time. The dependent variable, theme park visitation numbers, was taken from a reputable source, The Park Database website, which is known for its comprehensive data on theme parks. This primary variable is critical throughout the research to describe the relationship animated content has on visitation rates.

In addition to theme park visitation data, various demographic control variables that would account for any variance in the data were also added. This variable included GDP, country population, and tourism rates, all taken from The World Bank database. By incorporating these variables, the analysis aimed to outline the animation industry's specific impact on tourism activity by isolating the distinct contributions different demographic factors of a region might enact.

I ran multiple panel data regression models with my data spanning eighteen years, seven countries, and nineteen theme parks. This method facilitated the understanding of tourism data within different countries over almost two decades, which proved helpful in proving the effects the animation industry has on economic activity. My main findings underscore a significant correlation between the animation industry's popularity and increased attendance at theme parks. Thus, the growth of the animation industry plays a role in economic development.

The rest of this research is organized as follows: chapter two presents a critical review of five papers relevant to this research, and detail my theory in chapter three. I then discuss my data and methods in chapter four, present my results in chapter five. Finally, I provide a discussion of results and conclusion of my study in chapter six.

CHAPTER 2

Literature Review

In this chapter, I analyze five empirical research papers on the relationships between the animation industry, economic activity, and film-induced tourism. The common message across these articles is that they talk about film-induced tourism as a phenomenon where individuals become inspired by places they see on screen and travel to those locations and experience them firsthand; this phenomenon has emerged as an essential driving factor for tourism today. Moreover, the articles discuss that animation popularity has played a major role in driving theme park visitations numbers as more and more people are becoming attracted to the idea of visiting theme parks to "meet" their beloved characters in-person. Through analyzing these research papers, I set context for my own contribution to this subject.

2.1 The Power of Anime: A New Driver of Volunteer Tourism by Hiroaki Mori

This research paper analyzes the connection between anime-induced tourism and volunteerism, a form of tourism in which individuals are stimulated to tour a specific country based on the animation-related content they were exposed to. The article primarily highlights anime enthusiasts' unique motivations and activities they indulge in by utilizing qualitative data, which includes the analysis of three case studies that were chosen based on criteria like "(1) the different types of organizers of volunteer projects; (2) the different geographical areas involved; and (3) the different storylines of the anime involved." Using qualitative research methods, the author analyzes various evidence sources, ensuring a well-rounded perspective. Moreover, the interviews with individuals deeply involved in anime tourism provide additional data, capturing nuanced insights and personal experiences.

The article heavily explores existing literature and other research papers on film-induced tourism and anime-related tourism to identify the intersection of these two components. Moreover, to better understand film anime-induced tourism, the paper implies methods including collecting data from various sources such as survey interviews and participant observations and documenting their behavior, focusing on tourists visiting anime-related sites and documenting their activity. The data collected for this research went through several empirical and statistical review processes, analyzing different trends and patterns on how individuals stimulated by different anime show interest in tourism in those anime-related sites. On top of that, the article talks about contrasting anime-related tourism with the more commonly known film-induced tourism and voluntourism (voluntary tourism). All of this is to understand how anime-related tourism contracts with the other stimuli people use to visit and tour different places related to one film medium or another. Additionally, in the analysis, the paper mentions how it is essential to analyze the impact on host communities and how the sudden influx of tourists affects the economy, how local communities perceive it, and what benefits or drawbacks it brings. All of these paper's mentions, although they have some connotations in the long run, do not negatively impact the hosting communities.

The analysis presented in this article indicates that anime fans have deep emotional ties with the locations associated with their favorite animes, which translates into a sense of "gratitude" toward the communities depicted in their favorite animated series. This leads to their active participation in clean-up activities at anime-inspired locations, and they are motivated to explore the country they are visiting in more detail. With that being said, the study mentioned environmental preservation as a predominant theme, showcasing the dedication to maintaining the integrity and beauty of the destinations that the participants of such anime conventions exhibit. The research underlines the positive interactions between anime tourists and host communities. This positive behavior fosters mutually beneficial relationships, creating a symbiotic relationship between visiting individuals and the local communities. The study notes that instances of negative behavior among fans are rare, emphasizing their genuine respect for the places they visit.

In conclusion, the study presented in the article regarding the relationship between anime-induced tourism and voluntourism employs a qualitative approach, incorporating various data sources such as case studies and interviews, to name a few. This methodology explains anime-induced tourism's motivations, activities, and outcomes; the study adds knowledge of this unique tourism phenomenon and highlights the potential for fandom-driven tourism to impact communities positively. By emphasizing anime enthusiasts' passion for the field, the research suggests a promising way to foster responsible tourism practices and establish genuine connections between tourists and host destinations.

2.2 Cartoon Planet: Worlds of Production and Global Production Networks in the Animation Industry by Hyejin Yoon and Malecki Edward J.

The following article examines the highly complex and ever-evolving global dynamics of the animation industry. The authors Hyun Bang Yoon and Edward J. Malecki look in-depth into the many steps that go into creating animated content, examining all the ways, starting from traditional 2D animation to modern computer-generated imagery (CGI).

Hyun Bang Yoon and Edward J. Malecki thoroughly analyze the five stages of animation production in the article. These steps include conceptualization, pre-production, production, post-production, and distribution. The authors highlight the importance of the conceptualization phase because it is the "creative" phase. It involves character and background development through storyboards to visually convey the script's essence. The authors highlight the division of labor, where creative aspects like character development and scriptwriting are primarily managed in North America and Europe. At the same time, Asia handles labor-intensive tasks such as repetitive drawing and coloring. In addition, they emphasize the impact of globalization, explaining how animation studios worldwide collaborate in an organized manner, where they efficiently allocate tasks based on cost and location and other details that go into the production of animated content.

The article also looks into the animation industry's shift from traditional 2D to modern 3D animation. The authors emphasize the high entry barriers and capital requirements of 3D animation, which explains why only significant studios can afford such complex projects. They also illustrate emerging studios' challenges, including the risks associated with a fluctuating market and the many resources needed for

CGI production. The authors emphasize the industry's adaptability, showcasing how studios navigate between markets and platforms to seize opportunities.

Another interesting fact the article touches upon is the term "worlds of production," which the author explains within the animation industry. This term reflects a separation from the traditional film production process, which is governed by the unique technologies and labor input utilized in creating animated content. The productions of animated film content are woven into the global production network quite intricately. At the same time, they showcase a distinct form of globalization compared to traditional media outputs; the author also mentions that, unlike conventional film productions, the animation industry is characterized by globalization and international divisions of labor mainly because this field requires a large pool of artistic talent, which creates a high volume of labor demand even across borders. Overall, the article touches upon how incredibly demanding this industry has grown to be because, as mentioned previously, it requires more and more labor with every growth, meaning that this industry creates many new job prospects and thus, as the article implies, can be a substantial factor in stimulating the economy through the increase in job market demand.

In conclusion, the article explores the animation industry's global landscape in-depth. Its examination of the production stages, industry strategies, and international collaboration provides readers with an apparent understanding of this dynamic, ever evolving, and economically significant field. Nevertheless, obtaining more specific publication details is crucial to assess the article's scholarly context thoroughly. Additionally, while the article provides a comprehensive overview, a more critical analysis of the discussed trends and their potential implications for the animation industry's future would enhance its depth.

2.3 The Prevalence of Storyworlds and Thematic Landscapes in Global Theme Parks by Carissa Baker

The article starts by saying that the global park industry has experienced significant growth and transformation over the years and emphasizes that one of the attributes of this transformation is that the parks started to employ storytelling to enhance visitors' experiences. In the article, Clarissa Baker, who serves as an assistant professor of park attractions and management, delves heavily into the empirical research of this industry to outline any existing patterns that might expand its ever-growing popularity. The paper thoroughly analyzes this industry, enacts different methodologies, and produces empirical results.

In this paper, Baker focuses on two significant aspects of the parks - theming and storytelling in the global theme park industry. The article mainly distinguishes parks from amusement parks and highlights that the addition of "thematics" to parks is the driving factor that attracts many individuals, which sheds light on the significance of themed elements and the aspect of storytelling in the parks that creates an immersive environment driving many individuals wanting to participate in it. The paper also mentions the regional variation in the park narrative employed, which sheds light on the cultural aspect of theme parks and its influence on visitors' experiences.

The article employs empirical analysis by drawing data from various industry reports and other scholarly reports. To be more specific, Professor Baker uses a quantitative approach to analyze and look deeper into the presence of thematic elements, narrative-driven amusement rides, and many storytelling strategies across different theme parks. To be more specific, the study utilizes descriptive statistics, such as describing what percentage of parks have incorporated three or more themed land or how common themes are distributed among parks. Thus, the author emphasizes that top-tier parks have a typical pattern of employing thematic elements, highlighting their attempt to create immersive environments and attracting more individuals to the park this way. Moreover, Baker also talks about how different parks sometimes have a common theme based on the region or culture in which they are located. This also outlines the importance of culture on theme park thematics and thus its visitation rates.

The empirical analysis professor utilizes data from multiple sources, which, as mentioned above, are industry reports, theme park-specific information, and various other research papers and scholarly reports. With that being said, Baker heavily utilizes theme park visitation statistics to determine the most "popular" or heavily attended theme parks globally. Additionally, the paper draws upon data on various other factors that contribute to theme park visitations, such as tourism rates and the narrative theory, to contextualize empirical results and the overall importance of thematic representation in the popularity of the parks. Additionally, the article uses several case studies to back up the findings mentioned in the empirical analysis. One example is SeaWorld Park, which is located in North America. Baker mentions the park, along with Ocean Park and Yokohama Sea Paradise, empathizing their seal life thematic override narratives and thus have a certain sense of popularity due to that.

Overall, the following research appears to provide valuable insight by producing literature that showcases the various thematic factors that contribute to the overall popularity of various theme parks both within the borders and beyond them. The paper uses empirical results, as mentioned, to outline various descriptive statistics to the reader in order to underscore the importance of thematic narratives along with narrative experiences in drawing visitors to these parks. Additionally, the article touches upon briefly some managerial implications with the incorporation of a thematic aspect into a park, which, as the author mentions, is important for ensuring the park's competitive position in the market. Overall, the empirical analysis presented provides valuable insight into the ever-evolving theme park landscape and sets up the way for future research papers interested in analyzing the importance of a theme aspect in park popularity both on a local and global scale.

2.4 Film-induced tourism– the impact of animation, cartoon, superhero, and fantasy movies by Lazaro Florido-Benítez

Film-induced tourism and its implications on theme park visitation became crucial areas of study in the field of entertainment and tourism, as outlined by Prof. Florido-Benítez. The following research paper explores the dynamic relationship between different types of films, parks, and tourism rates. The primary focus is on understanding how movies, especially animation, cartoons, and superheroes from fantasy genres, contribute to the visitor experiences in a theme park. The study emphasizes the influence of major players like Disney and Universal Studios, shedding light on how the popularity of film content produced by these two companies ensures the popularity associated with theme parks for these companies.

The foundation of this research paper is based on the data that showcases the top 20 themed parks globally by the number of visitors from 2010 to 2021. The data is delivered by Themed Entertainment Association (TEA/AECOM) and The Number of Nash Information Services LLC. The data provides the ranking of the themed park, its location, and the number of total visitors from 2010 to 2021. This research paper touches upon a wide range of geographic locations, including major theme parks in the United States, Japan, France, South Korea, Germany, and Hong Kong.

The research paper employs a multi-faceted approach to analyzing the correlation between film-induced tourism and themed parks. The study focuses on the strategies implemented by major companies, specifically Disney and Universal Studios, to leverage movies as a marketing tool. Through a detailed examination of attraction, the paper explores how these themes extend movie experiences to physical spaces. The analysis delves into the significance of creating an immersive environment based on the film's popularity and various characters. The paper also touches upon the impact of external factors, such as the COVID-19 pandemic, on the film and theme park industries. It emphasizes the resilience of Disney and Universal Studios during the pandemic, noting their ability to adapt to local travel restrictions and maintain visitors. The research also goes beyond immediate data, acknowledging the potential future shifts in the global theme park landscape, particularly in Japan, driven by the popularity of manga characters.

More specifically, the paper's main goal is to investigate the impact of film-induced tourism on park visitation rates by analyzing the importance of popular movies associated with different theme parks starting from the year 2010 to 2022. By utilizing both qualitative and quantitative data, the author sets up the research by first defining the most heavily visited or popular theme parks and then identifying the most watched animations and movies that are in one way or the other linked to the theme parks. The overall key findings reveal that theme parks that include popular and fairly recognizable characters such as Harry Potter, minions, and various Disney characters tend to be more popular by capturing the interests of wider audiences. The author mentions that these characters are incorporated into various themed attractions and thus provide a fantasy world experience to visitors, which is the driving lure of such destinations. The paper also outlines a highly strong correlation between film-induced tourism and theme park visitation: the author mentions that around 215 million people visited the top 20 theme parks worldwide. This, as the author mentions, is a significant stimulus for local economies mainly due to the fact that such activity drives additional revenue to that region. Overall, the article says that these parks serve as an extension to various media experiences, facilitating a transition from movie screens to real-world environments and enriching the overall visitor experiences.

In conclusion, the paper provides a comprehensive overview of the correlation between the animation industry and tourism rates, highlighting the integral role of movies and other media content in attracting people to different places. The data provided by TEA/AECOM serves as a robust foundation for analysis, offering insight into global trends in theme park visitations. As the study suggests, theme parks' success relies on their ability to create a seamless transition from cinematic worlds to physical attractions, creating a continuous and captivating journey for visitors, which will ensure high tourism-related visitation rates. Overall, the paper concludes that the animation industry can be a valuable tool to market different locations and inspire people to visit them by attracting them to theme parks associated with famous film content such as movies, cartoons, or Superheroes.

2.5 The Characteristics of Film Products to Induce Tourism by Noridawati Abd Rahman, Zairul Anwar Dawam, and Jennifer Kim Lian Chan

The article investigates the relationship between film products and their influence on tourism, dissecting various characteristics that contribute to the attraction of tourists to potential locations. The article touches upon six aspects that govern this relationship including on-location filming, representation of modern architectures, cultural elements, authenticity of filming locations, the impact of storytelling, and the role of celebrities all of which provides a holistic dynamic that plays a role in promoting film tourism. By addressing these diverse factors, the paper aims to broaden our understanding of how film acts as a potential stimulus for tourism and shape potential travels' perception of destinations.

One strong aspect of the article lies in how extensively it analyzes the literature cited, which sheds light on the study within the existing knowledge on film-induced tourism. The addition of various works to the study, such as the studies emphasizing location filming and the debate on authenticity versus substitution of filming locations, add particular depth to the overall discussion. The author effectively navigates through the diverse literature, highlighting the film industry's and tourism's interconnectedness while acknowledging the nuanced perspectives surrounding this relationship.

As for the methodologies the paper implements, it demonstrates a robust approach to understanding the impact of film-related products on tourism. The study showcases comprehensive research through analyzing data from case studies, surveys, interviews, and content analysis of various films. The mixed methods approach allows for a nuanced exploration of complex interplay between film and tourism, which ideally captures both qualitative insights into tourists motivation and perceptions and quantitative measurements of tangible impacts on visitor arrivals and economic indicators.

The paper, as already mentioned, has the goal of investigating the relationship between various film-related products and tourism rates. The aim is to understand how various movies and television shows inspire people to visit otherwise very distant lands and locations. By analyzing the six characteristics of film products, the author examines the ways in which these factors contribute to film tourism. However, the study also mentions that despite the potential for film-induced tourism to drastically increase visitor numbers, there are certain complexities associated with the difficulty of gathering empirical data and the overall complex nature of the industry. The author mainly says that the success of a film or TV franchise does not guarantee the popularity of a location. Nevertheless, analyzing these factors and continuing research on their impact is important to enhance the tourism industry and attract more visitors.

The article provides an in-depth exploration of the impact of film on motivating individuals to visit a particular country. However, given the ever-evolving nature of the film and tourism industries, there is room for improvement, mainly by incorporating and analyzing the impact of recent technological and social advancements on how people participate in film-induced tourism. In addition to addressing potential limitations of the methodology of this paper, it is also important to outline what possible contributions this article could provide to the overall research of this topic. Nonetheless, the study serves as an interesting medium for discussing the interplay of film in promoting tourism.

2.6 Summary

To summarize all five articles, it is safe to say that all of the mentioned articles acknowledge the ever-growing significance and popularity of the animation industry along with the influence of this newly established phenomenon of film-induced tourism that came into play with the rise of digital platforms. All the articles also identified a similar pattern: as modern technologies became more and more popular, the animation industry grew alongside it, not to mention that this growth transcends borders and reaches global audiences. Additionally, the article mentions that film-induced tourism is growing rapidly and attracting more individuals to visit familiar places and characters from their favorite well-known animated films. This allure of visual incentives comes in different forms for individuals, as some want to experience the geographical locations seen in their favorite cartoons, and others want to "meet" their favorite cartoon characters in person.

Additionally, the articles all underscore that individuals are attracted to visiting geographic locations seen on screen, irrespective of their geographic location. This phenomenon, as mentioned, brings to light the power of storytelling and how fictional scenarios can inspire real-world experiences. As a result, the destinations featured in animated films or places depicting famous characters tend to benefit quite significantly economically by attracting tourists to that region. This symbiotic relationship between the animation industry and tourism rates highlights the potentially transformative power the animation industry possesses on local economies.

Furthermore, all five articles mention that due to the transformative power of the animation industry, different locations can capitalize on the popularity of animated productions to leverage that popularity and attract more and more tourists to a specific area. Thus, it was concluded that the animation industry is becoming a massive force shaping the behavior of tourists and influencing their decisions, which highlights the need to study this field and utilize it as a possible marketing tool.

Overall, the animation industry possesses immense potential as a factor attracting the interest of tourists and fostering the growth of local economies in targeted or hosting countries. Overall, This underscores the need for continuous research in this newly emerged field. Future researchers could continue leveraging the preexisting literature to dive further into the symbiotic relationship that governs the impact the animation industry has on tourists’ choices.

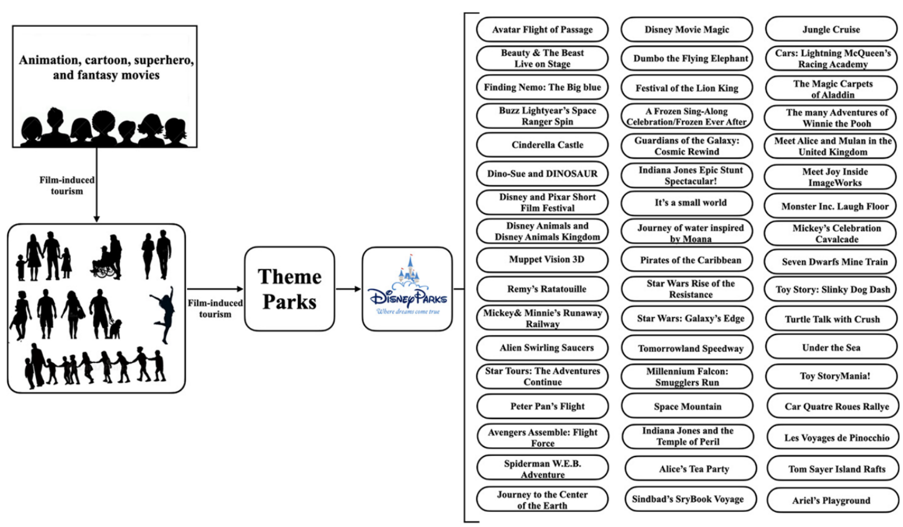
CHAPTER 3

Theory

The following theory chapter is structured in two parts to analyze the animation industry both on micro and macro levels. However, the theory will initially use the Disney and Universal models adapted from Prof Florido-Benítez's work to conceptualize the theory framework. At a macro level, the theory will look closely into how the popularity of different popular animated content affects the utility individuals get from visiting cities related to animation either by direct geographic portrayal or by showcasing well-known characters. Additionally, a follow-up analysis will also be drawn to see how the demand for theme park visitations is affected by it. The analysis on a micro level is targeted to understand individuals' preferences and behaviors in response to fluctuations in the popularity of various animations. On the other hand, the analysis of the theory on a macro level will explore the overall broader influence of animation popularity on local economies. To be more specific, the theory will show how the increase in the popularity of the animation industry as a whole affects consumption and investment factors that go into the aggregate demand curve, which leads to a net positive effect on employment artists and the overall country's GDP. Thus, by analyzing the animation industry both on a micro and macro level, this chapter attempts to deliver a detailed explanation of the effects the animation industry has on the economy.

3.1 Adapting Professor Lázaro Florido-Benítez's Graphs

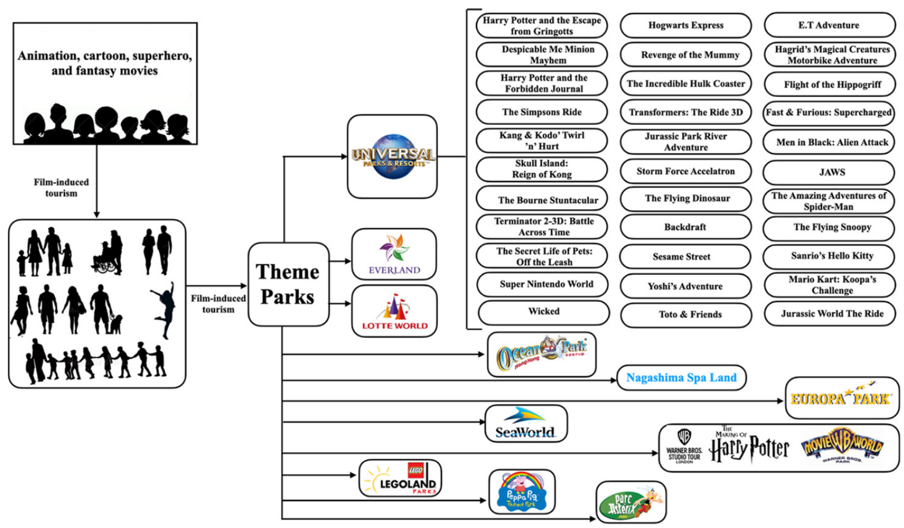
To further explain the theory, a decision was made to incorporate graphs created by Professor Lázaro Florido-Benítez. More specifically, his models offer valuable insight into how popular cartoons stimulate individuals to participate in film-induced tourism in order to explore the locations portrayed in that cartoon or see well-known characters portrayed. Primarily, the Disney research model, seen in Figure 3.1, shows that viewers are enticed to visit locations depicted in cartoons as well as "meet" famous characters showcased in cartoons like "Frozen" or "Beauty and the Beast." This emotional connection creates a desire to experience these fictional settings firsthand, which results in an increased number of visitations to destinations that mirror these iconic places and characters. Theme parks such as the ones created by Disney serve as a tangible manifestation of fictional realities, providing visitors with the opportunity to interact with fictional characters and immerse themselves in various captivating stories.



Source: Own elaboration

Figure 3.1: Disney's Research Model (Source: Film-Induced Tourism in Animation, Cartoons, Superheroes, and Fantasy Worlds)

Similarly, the Universal research model depicted in Figure 3.2 shows how popular cartoon characters featured in various theme parks created or spanning from Universal Studios attract individuals, which also results in people participating in film-induced tourism. Famous products created by Warner Bros., as well as the world of Harry Potter, provoke a sense of wonder and excitement and thus encourage visitors to visit those theme parks and attractions that actively promote well-known characters. Not to mention Universal Studios' immersive experiences allow visitors to step into the shows of their beloved characters, which is the driving force for film-induced tourism.



Source: Own elaboration

Figure 3.2: Universal's Research Model (Source: Film-Induced Tourism in Animation, Cartoons, Superheroes, and Fantasy Worlds)

Overall, these models lead to the assumption that the implications of film-induced economic development can be potentially significant. By adopting the appeal and lure of popular cartoons and iconic characters, various geographical destinations can enhance their attractiveness and possibly even differentiate themselves in a competitive tourism market. Moreover, the influx of tourists driven by film induced tourism increases the revenue in local economies, which supports businesses ranging from hospitality all the way to entertainment industries. Additionally, investing in the park attractions further stimulates economic activity and job creation, which contributes to sustainable growth and development. Overall, Professor Lázaro Florido-Benítez's research models shed light on the interconnected nature of animation and tourism, which highlights the power of popular cartoons to shape travel habits and behaviors.

3.2 Micro Level Analysis

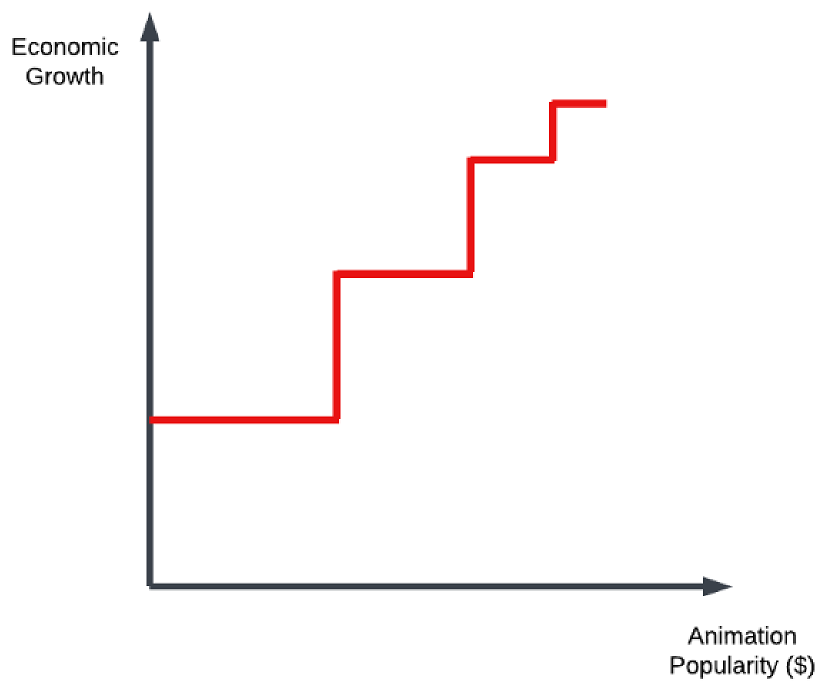


Figure 3.3: Model Showcasing the Economic Impact of Animation Popularity

Animations can captivate audiences worldwide, which is the foundation of the relationships between the animation industry and economic development. Audiences are attracted to various animations due to their ability to create relatable characters and fictionally desirable settings, which drives them toward film-induced tourism when they see these creations on screen. Moreover, attractions such as

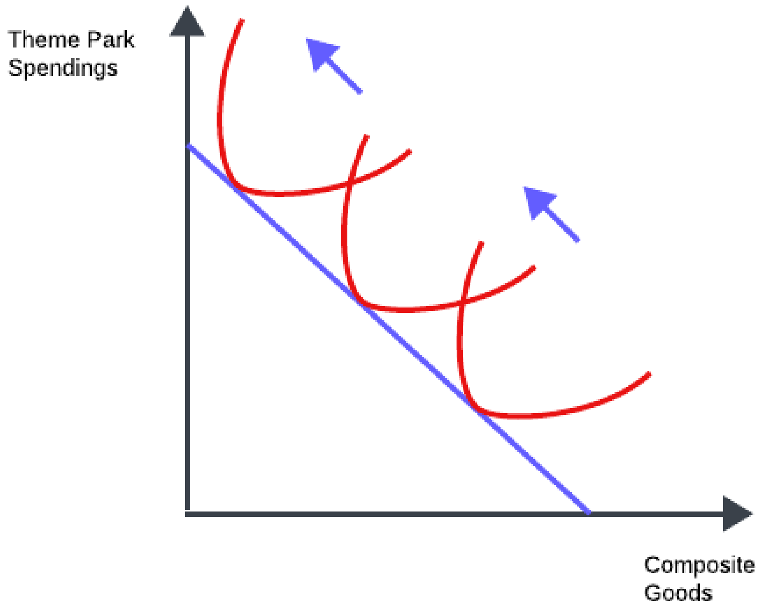


Figure 3.4: Model Showcasing the Utility Maximizing Curve in the Animation Industry

themed parks, comic events, and anime conventions increase tourism rates as well as strengthen the local economy through the capital brought in by ticket sales, merchandise sales, and hospitality services. The following is due to the established curiosity to see these imaginary worlds in person. This relationship is based on the established emotional bond between the traveler and the animated narrative, influencing travel trends and stimulating economic activity.

To better describe this relationship, as seen in Figure 3.3, there are two leading components of this symbiotic relationship - the revenue generated by film tourism sites and the popularity of different animations, which is measured through the profits they generate, meaning the higher the revenue, the more popular the content. As seen in this model, a substantial increase in the popularity of animated content, given that a cultural representation takes place in that particular content, increases the number of visitations to a specific site of tourism, which is evident with the sudden increase in revenue. That being said, this relationship is not predicted to hold constant as animation's popularity increases over time. Its effects on the visitation rates of tourism sites will lessen with every increase, reaching a state at some point where the relationship becomes perfectly inelastic.

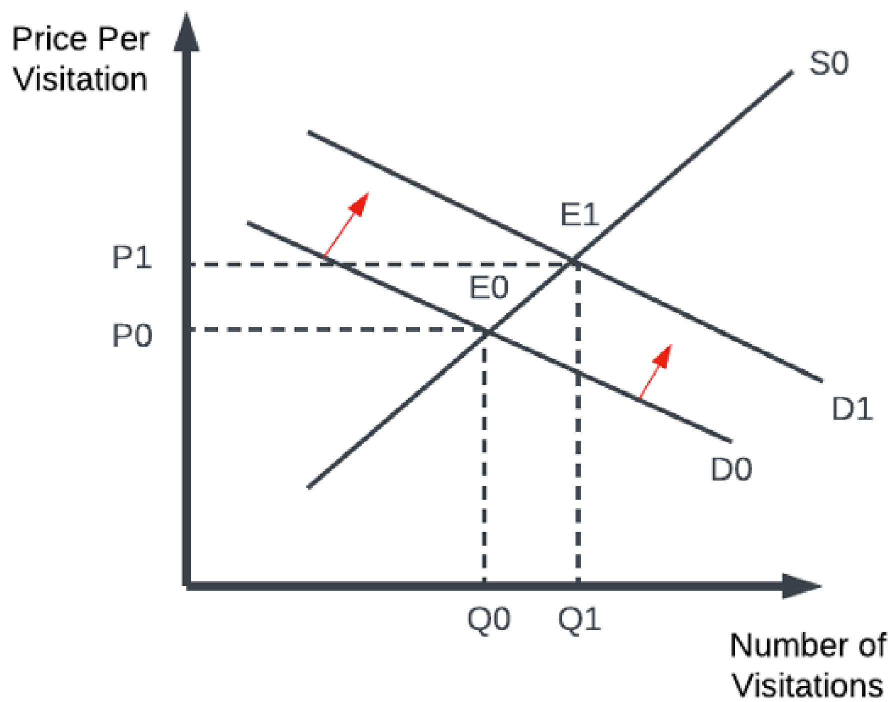


Figure 3.5: Market For Film Tourism

As seen in Figure 3.4, the utility maximization curve portrays the relationship between spending associated with the animation industry and related products on the y-axis and composite goods on the x-axis. These graphs illustrated that with the popularity of different animated films, individuals gain more utility by spending more of their disposable income on products associated with these animations. This is shown by the upward shifts along the utility maximization curve. This suggests that people are willing to allocate more of their spending towards the animation industry and its associates, and this additionally gives more units of satisfaction or utility for each individual. Overall, the utility maximization curve shows how the changes in animation popularity influence consumers' spending habits and their performance, which is, in the grand scheme of things, affecting one's overall satisfaction levels.

The net effect of animation popularity should be examined to develop this model further. As seen in Figure 3.5, the increase in the popularity of animation will theoretically increase film-induced tourism rates, consequently affecting the number of visitations to a particular tourist site. This effect will cause the demand curve to shift to the left, increasing the price of visitations, which is defined as the prices for fairs and entry tickets, etc.

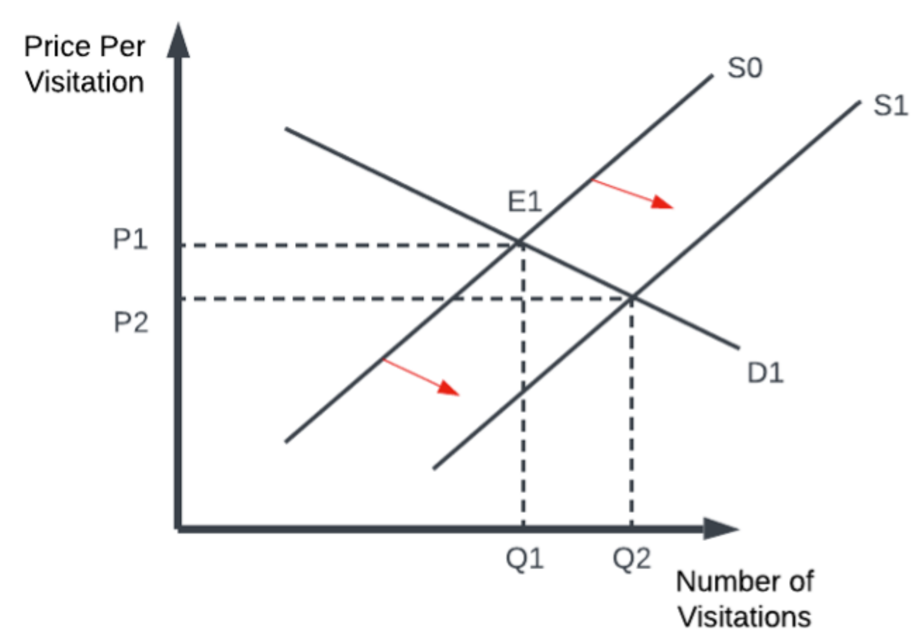


Figure 3.6: Market For Film Tourism

Additionally, as seen in Figure 3.6, the increased number of demands will cause the price of visitations to increase, as seen by the right shifting of the supply curve to meet the newly established demand for visitations. This will result in new equilibrium price and quantity levels, which will consequently benefit the hosting country's economy.

3.3 Macro Level Analysis

The Aggregate Demand (AD) and Supply (AS) framework is an essential tool in macroeconomic analysis, which provides an overview of how the economy is performing overall. Aggregate demand represents the total quantity of goods and services, which sectors within the economy are willing to and/or capable of purchasing at a specific price level and at a specific time. AD considers consumer spending (C), investment (I), government expenditures (G), and net exports (X-M) or, in other words AD = C + I + G + (X-M). Whereas aggregate supply represents a framework showcasing how much goods and services all the producers within an economy are capable or willing to supply at a given price level at a certain time.

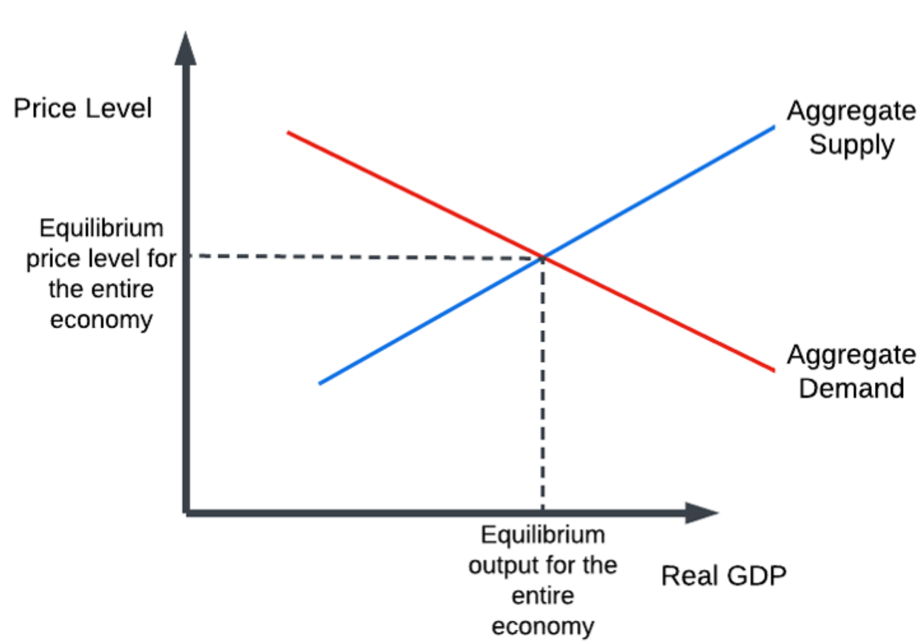


Figure 3.7: Aggregate Demand and Supply Curves

The intersection of the AD and AS curves, as seen in Figure 3.7, represents the equilibrium price level and output of an economy. This framework is an essential tool used by many economists and policymakers to analyze and address issues such as inflation, unemployment, and economic growth through the process of identifying the factors that influence the overall activity within the economy.

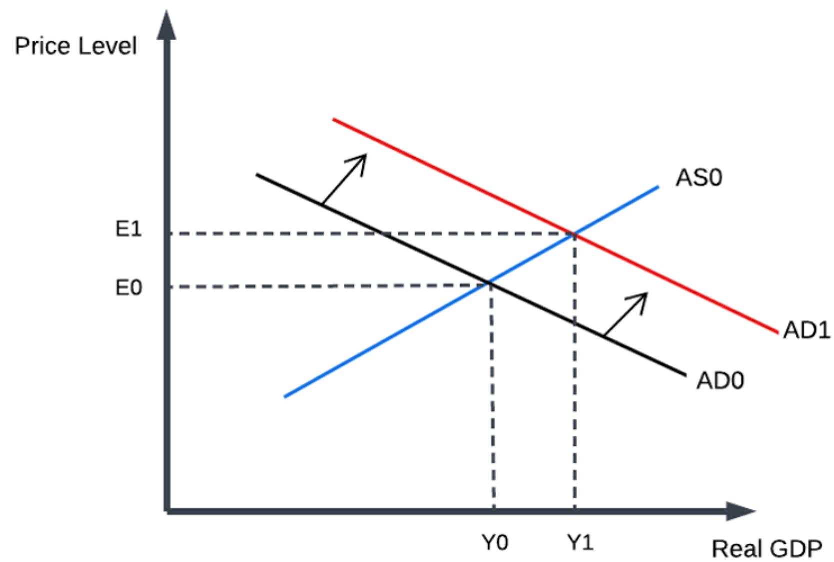


Figure 3.8: The Shift of the AD Curve to the Right

The exposure of the economy to the animation industry causes notable shifts in both the aggregate demand and supply curves, which contributes to the overall beneficial economic growth. Primarily, the demand side of the curve increases due to the rising rates of consumption (C), caused by individuals captivated by this industry. Such individuals increase their spending on related products and services, not to mention due to the increase in tourism rates caused by the animation industry, related services such as restaurants, hotels, souvenir shops, etc. also see a drastic increase in profits. Moreover, investments (I) also increase as the animation industry often requires sufficient capital inputs for production and technological advancements. As for exports, with the increase in the popularity of the animation industry internationally, exports (X) might increase as well causing a positive increase in net exports (X-M). Government spending (G) might be influenced indirectly through policies that support the creative and entertainment sectors.

Therefore, as seen in Figure 3.8, due to the expansion of almost every component that comprises aggregate demand, this curve will shift to the right.

Consequently, the aggregate supply will adjust to the increase in the components of the AD curve. More specifically, investments towards the technological advancements of this industry can potentially increase efficiency and productivity, leading to a positive impact on the AS curve, not to mention that the growth of this industry might also target the promotion of skilled labor, which would further benefit the aggregate supply curve. The AS curve will shift rightwards in order to respond to the increase in the AD curve, resulting in a new stable equilibrium output. Overall, the economy will experience beneficial growth from the exposure to the animation industry, which, from the model above, showcases the interconnected nature of different economic components, primarily with consumption, investments, and exports playing a crucial role in shaping the economic landscape.

3.4 Results and Predictions

The animation industry, in general, has experienced exponential growth over the past several years, and this trend is expected to continue, which is reflected in the growth model presented. People all around the world are starting to "consume" more and more animated products, which benefits the industry and increases popularity by the day. As a result of this cycle, economies worldwide increasingly utilize animations for marketing their products and thus reach publicity, increasing demand for animated content from animators even more. This, in turn, leads to more economic activity, which impacts the growth of a country.

Not to mention the animation industry will also be transformed by emerging technologies, as this aspect of the industry is ever evolving and developing. The growth of this industry can be attributed to many factors, such as modern technological advancements and the rates of globalization. Moreover, the advances in digital technology have made animation production more efficient and profit-maximizing, thus driving growth in the industry. As expected, this industry will positively impact a country's economic output and overall well-being. Therefore, these innovations are projected to revolutionize how animations are created and consumed; artists will be able to take advantage of advanced tools provided by software and hardware advancements to produce more realistic and detailed animations. Virtual reality technology may offer additional ways for animators to engage their viewers. As a result, this industry is expected to see an increase in demand, popularity, and consequently output.

In conclusion, while the animation industry is expected to grow in the coming years, economic predictions depend on various factors that are difficult to predict with certainty. It is crucial to keep an eye on market trends, consumer behavior, and technological advancements to stay up to date with the latest developments in the industry. Nevertheless, the current results and output of this industry are still not to be undermined as this industry generates substantial income annually, along with producing many positions for potential employment.

CHAPTER 4

Data and Methods

The paper focuses on understanding how the animation industry impacts economic development by increasing the tourist activity of the portrayed region. In other words, the research aims to prove that large animation companies like Disney, Pixar, Dream Works, and Universal Studios can be used as marketing tools by either portraying cultures and regions within animations or utilizing famous, well-known characters in designated theme parks to stimulate people to visit and explore specific location.

Given the unique nature of the research topic chosen and the lack of pre-existing datasets, the decision was made to create a custom-designed dataset, which would underscore the need for a novel approach. The subsequent sections explain the process employed in building the dataset with the **AnimationPopularity** variable as the central variable of interest along with other essential variables such as **Country**, **Year**, **Park**, **Visitations**, and several control variables such as **Tourism**, **GDP**, and **Population**.

4.1 "AnimationPopularity" Variable

The initial phase of creating the dataset for this research involved carefully selecting three search terms – namely "Animated Film," "Animated Movies," and "Animated Series." These terms were chosen to provide a comprehensive view of the popularity of animations across different mediums and possibly develop an animation popularity index to determine the popularity of the animation industry in seven selected countries: USA, France, Germany, Netherlands, Japan, South Korea, and Hong Kong for the years starting from 2004 going up to 2022.

After collecting the raw search interest data from Google Trends, which is pulled from a random, unbiased sample of Google searches, it was necessary for the purposes of the methodology of this research to normalize the collected data. The data gathered from Google Trends was distributed by months across different years. Each search term ("Animated Film," "Animated Movies," and "Animated Series") was individually averaged throughout the months, which resulted in a singular number representing the Google search popularity of each term from 2004 to 2022. This step aimed to eliminate fluctuation across the data and provide a more coherent trend analysis. The data was further refined by averaging the three search terms collectively for each year, producing a comprehensive animation popularity index ("AnimationPopularity") for the specified timeframe. The process of averaging the data in two steps simplified the dataset and ensured that the subsequent analyses were based on robust and normalized metrics, allowing the accurate analysis of the animation popularity index trends across the selected period.

4.2 "Park," "Country," and "Year" Variables

The variables "Park," "Year," and "Country" were selected to ensure a comprehensive analysis of park attendance data and, at the same time, guarantee that the data was globally encompassing. The "Park" variable entrees were gathered from The Park Dataset database, which represents a reputable and reliable insight into the popularity and significance of various amusement parks worldwide. For this research, 19 amusement parks were selected globally based on their popularity. These amusement parks include 'Universal Studios Orlando,' 'Universal Studios Hollywood,' 'Magic Kingdom,' 'Epcot (Walt Disney World),' 'Disneyland Anaheim,' "Disney's California Adventure," "Disney's Animal Kingdom," 'Lotte World South Korea,' 'Everland South Korea,' 'Efteling Theme Park Resort the Netherlands,' 'Universal Studios Japan,' 'Tokyo DisneySea Japan,' 'Tokyo Disneyland Japan,' 'Nagashima Spa Land Japan,' 'Ocean Park Hong Kong S.A.R.,' 'Hong Kong Disneyland Hong Kong S.A.R.,' 'Europa-Park Germany,' 'Walt Disney Studios Park France,' and 'Disneyland Park France.' The following list was created based on the popularity of amusement parks worldwide, and as mentioned above, it ensures global representation by including parks from different regions, countries, and continents.

The "Year" variable was chosen, with the starting year being 2004, which aligns with the launch of the Google Trends platform. Meanwhile, the upper bound was set to be 2022, representing the most recent year for which The Park Database data was fully available. The selected time frame allowed for a comprehensive global amusement park popularity trend analysis across almost two decades. It captures the significant developments in the entertainment industry, mainly focusing on amusement parks, such as the rise of new technologies and evolving consumer preferences. The following adds robustness to the dataset as the panel data presented allows for deciphering long-term patterns while also considering recent influences that may have shaped the popularity of amusement parks worldwide.

The "AnimationPopularity" variable plays a vital role combined with the "Park" variable, grounded in the belief that the popularity and recognition of different animated movies that represent various characters of well-recognized animations contribute to attracting visitors to amusement parks. The relationship between the popularity of different animated films and associated amusement parks creates the ground for this research to develop and analyze the relationship between the two variables. Additionally, studying this relationship could enable this research to examine the extent to which different animations attract visitors from overseas to foreign countries.

4.3 "GDP," "Population," and "Tourism" Variables

The dataset also incorporates three control variables to account for the potential variance in the relationship between "AnimationPopularity" and "Parks" – namely "GDP," "Population," and "Tourism." These control variables are essential for ensuring the robustness and validity of this analysis, which allows us to detect if outside demographic or economic factors influence any observed trends in animation popularity and park attendance.

The data for these control variables was gathered from The World Bank database, a reliable data source on global economic and social indicators. Primarily, the tourism variable, defined by The World Bank as "International tourism, number of arrivals," offers insight into the influx of international tourism to a given country. More specifically, The World Bank defines this particular tourism indicator as the number of overnight visitors who travel to a country different from the one in which they have residency for a period that does not exceed 12 months. Thus, this variable encompasses tourists, same-day visitors, cruise passengers, and crew members.

By including these control variables, this research explores the dynamic relationship between animation popularity, amusement park popularity, and broader demographic and economic factors. The detailed sourcing of the data from The

World Bank ensures the reliability and consistency of the included control variables, which enhances the depth and validity of this analysis.

4.4 Data Cleaning

Combined, the dataset had a total of 51 missing values across the "Visitation" and "Tourism" variables. The gaps in the "Visitation" variable were due to the lack of data for certain parks in certain years in The Park Database from where the park visitation data was gathered. Whereas the "Tourism" variable exhibited missing values in all countries from the years 2021 and 2022 consistently. The data for this variable was gathered from The World Bank, and the lack of data can presumably be explained by the COVID-19 pandemic that started in 2020. Thus, to address these missing value discrepancies, a decision was made to impute the data by filling the missing entries with averaged-out values for individual parks. The average park visitation and country tourism rates were calculated for the individual parks across 18 years (2004 - 2022), and the missing values for those parks were filled accordingly. This approach aimed to mitigate the impact of missing values by providing representative estimators based on historical data, which ensured the integrity and completeness of the dataset used for this research.

4.5 Exploratory Data Analysis

Themed parks that showcase widely known characters have the potential to attract individuals from around the world; in other words, animation popularity and theme park visitation rates have a symbiotic relationship, which later translates into increased tourism rates when individuals from overseas are attracted to theme parks. Thus, for this research, 19 theme parks ('Universal Studios Orlando '

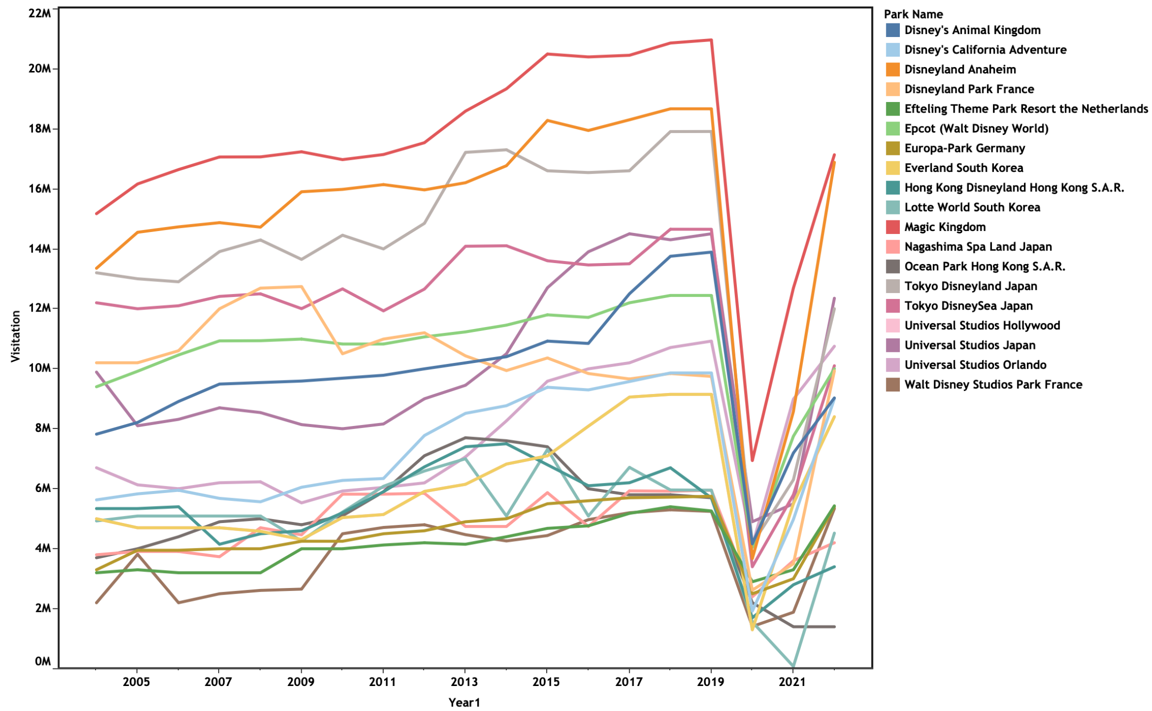


Figure 4.1: Visitation Trends from 2004 to 2022 for 19 Parks World-Wide

'Universal Studios Hollywood ' 'Magic Kingdom' 'Epcot (Walt Disney World) ' 'Disneyland Anaheim ' "Disney's California Adventure " "Disney's Animal Kingdom " 'Lotte World South Korea' 'Everland South Korea' 'Efteling Theme Park Resort the Netherlands' 'Universal Studios Japan' 'Tokyo DisneySea Japan' 'Tokyo Disneyland Japan' 'Nagashima Spa Land Japan' 'Ocean Park Hong Kong S.A.R.''Hong Kong Disneyland Hong Kong S.A.R.' 'Europa-Park Germany' 'Walt Disney Studios Park France' 'Disneyland Park France') were selected globally that had high visitation rates. As seen in Figure 4.1, there is a repeating pattern in the visitation rates of theme parks globally which we can see in the consistent upward trajectory from the year 2004 to 2022. However, it is also important to note that there is an anomaly present in the year 2020, where a sharp decline in global visitation rates can be clearly observed. This unanticipated dip in visitation rates can be explained by the COVID-19 pandemic and the safety procedures that were enacted due to it, which restricted people from attending public and massively crowded places.

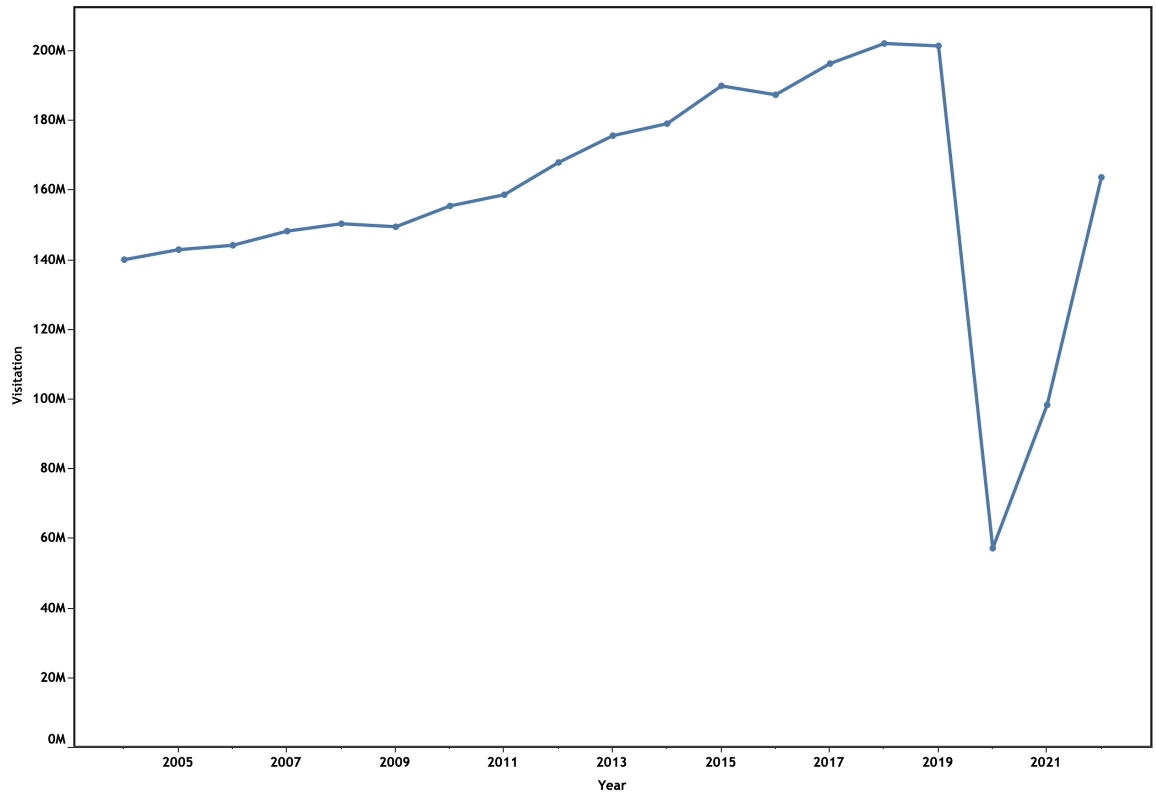


Figure 4.2: Mean Theme Park Visitation Rates Globally from 2004 to 2022

To further illuminate the growing popularity of theme parks, Figure 4.2 offers an overview of the average theme park visitation rates on a global scale. The graph showcases a distinct upward-sloping line that sheds light on this industry's consistent growth over the past 18 years (2004 - 2022). This could be used as evidence that emphasizes the growth of theme park popularity by featuring well-known, animated characters. This data also underscores the growing popularity of these parks as recreational sights attracting a growing number of individuals every year. Overall, the upward-sloping line showcases the expanding popularity of theme parks on a global scale.

Moreover, Figure 4.3 offers a more comprehensive view of theme park visitations by organizing the data by country. The graph offers a closer look at the distribution of theme park visitations on a country-by-country basis, which offers insight into the population trends across countries. With that being said, it can be seen on the graph that the USA, Japan, and France are the leading countries having the most amount of park visitors annually. Following those three countries are South Korea, Hong Kong, Germany, and the Netherlands, which also have a significant number of visitors in their parks every year. Overall, Figure 5.1.2 sheds light on the ranking of park visitations in various countries worldwide.

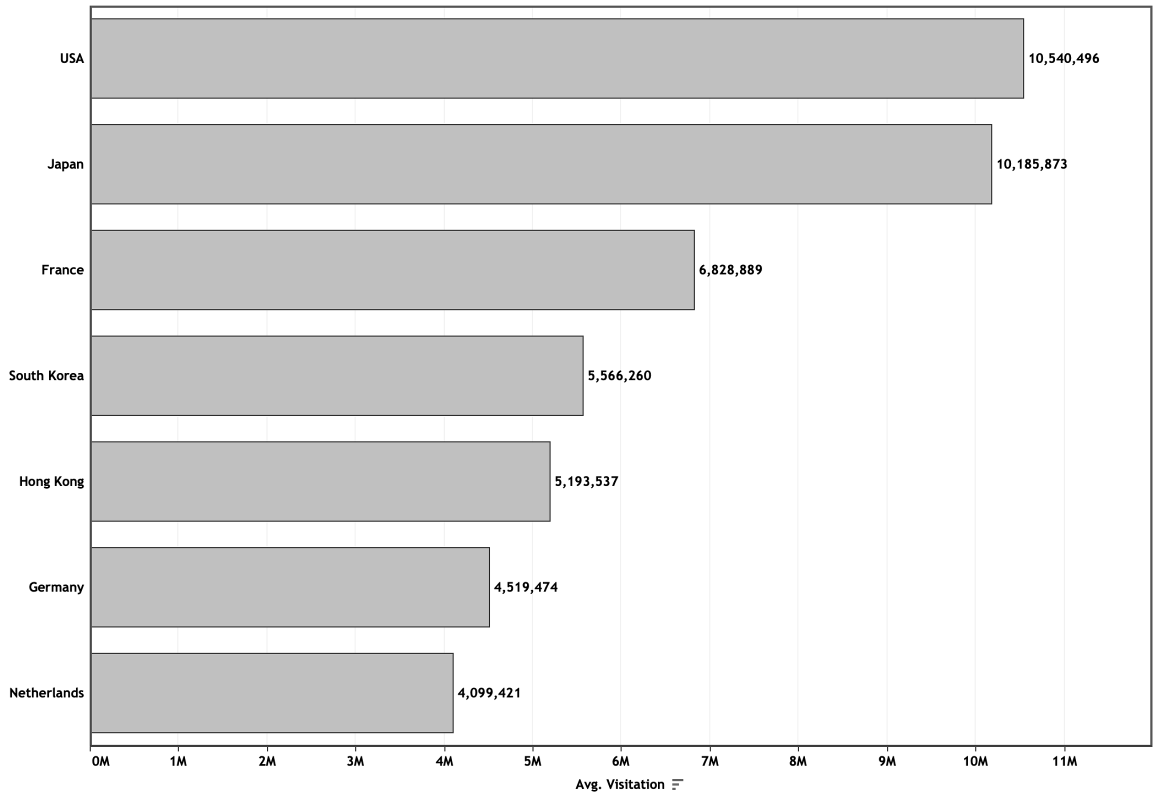


Figure 4.3: Theme Park Visitations by Country

It is also important to analyze tourism data in the selected countries. As mentioned in previous chapters, 7 countries were selected for this research ('USA' 'France' 'Netherlands 'Germany' 'South Korea' 'Hong Kong' and 'Japan'), and in Figure 4.4, it can be observed that France, USA, and Hong Kong are emerging forerunners that on average have consistently the highest number of tourists every year. Japan follows closely in fourth place which is followed by Germany in fifth place, then South Korea and lastly the Netherlands with the lowest number of tourists. Overall, the following visualization showcases conceptually the average number of tourists in the selected countries, which can later be used to evaluate whether a portion of these tourists are visiting for recreational purposes to visit the above-mentioned parks.

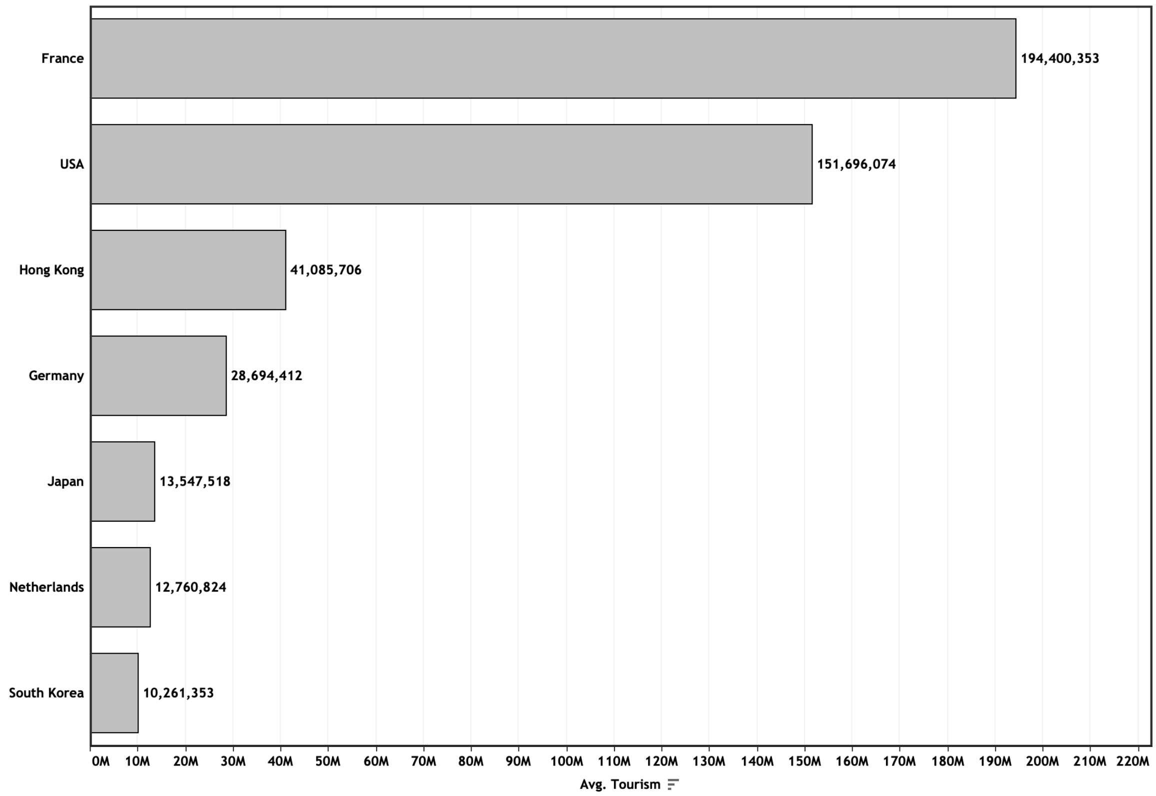


Figure 4.4: Total Tourism by Country

After rigorously examining tourism rates along with theme park visitation trends in various countries, the next step would be to explore the correlation between animation popularity, tourism rates, and park visitation trends on a country-to-country basis. The following is tied to the initial research question, which delves into whether the popularity of the animation industry increases economic development through stimulating tourism in particular regions. For this purpose, Figure 4.5 comes into play due to the fact that it looks over how these three variables interplay with each other. More specifically, it illustrates how both tourism and park visitation trends change as animation popularity increases. It can be observed that as animation popularity increases, park visitation increases in some instances, and with that, tourism rates also seem to increase. The following trend slightly differs based on the country and, in some instances, is difficult to identify due to the data being distributed by countries. Thus, while Figure 4.6 offers valuable insight into the matter, further analysis is necessary.

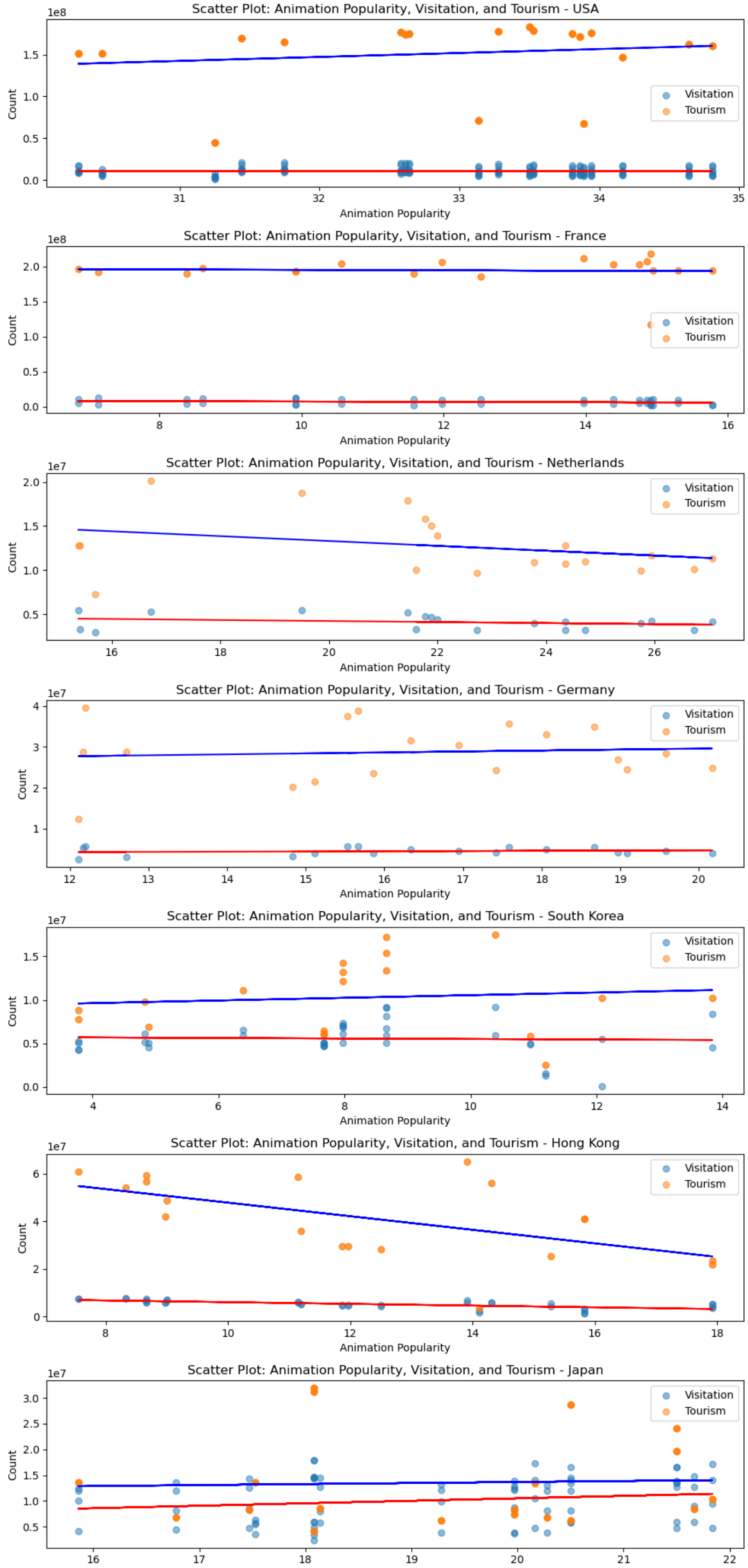


Figure 4.5: Animation Popularity, Visitation, and Tourism by Country

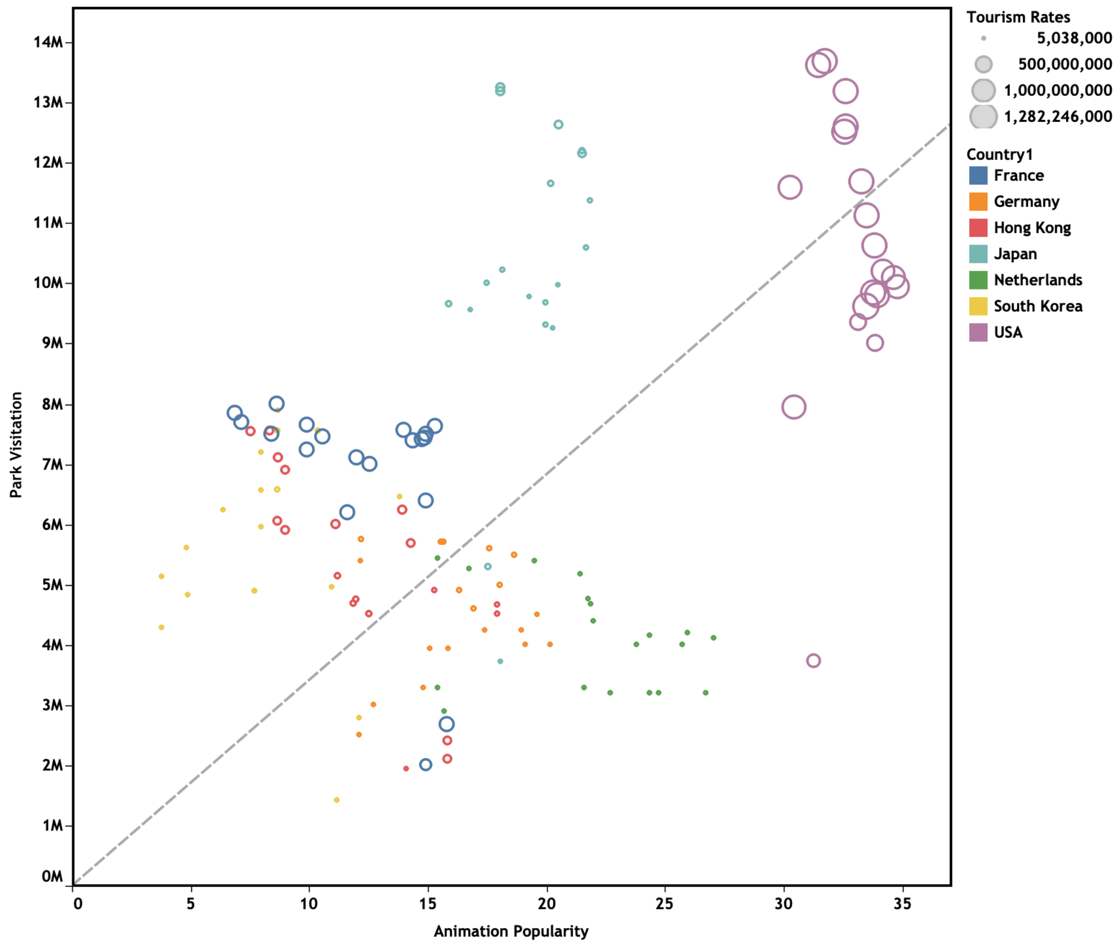


Figure 4.6: Animation Popularity, Visitation, and Tourism (y-Intercept forced to Zero)

To put the above-mentioned visualizations and the information drawn from them into perspective, Figures 4.6 and 4.7 reveal a significant level of correlation between animation popularity, tourism, and theme park visitation rate variables. Both graphs feature a line of best fit, with one having a y-intercept forced to go through zero while the other does not. More specifically, the presence of a strong correlation between the mentioned three variables is evident in the graph through the indicated statistical results. The R-squared value of 0.156328, quite a low standard error of 2.88151e+06, and a highly significant p-value (< 0.0001). These statistical measures suggest that there is a robust relationship between the variables. However, it is essential to mention that despite these findings, further investigation is required to test this relationship and understand the dynamics of these variables clearly.

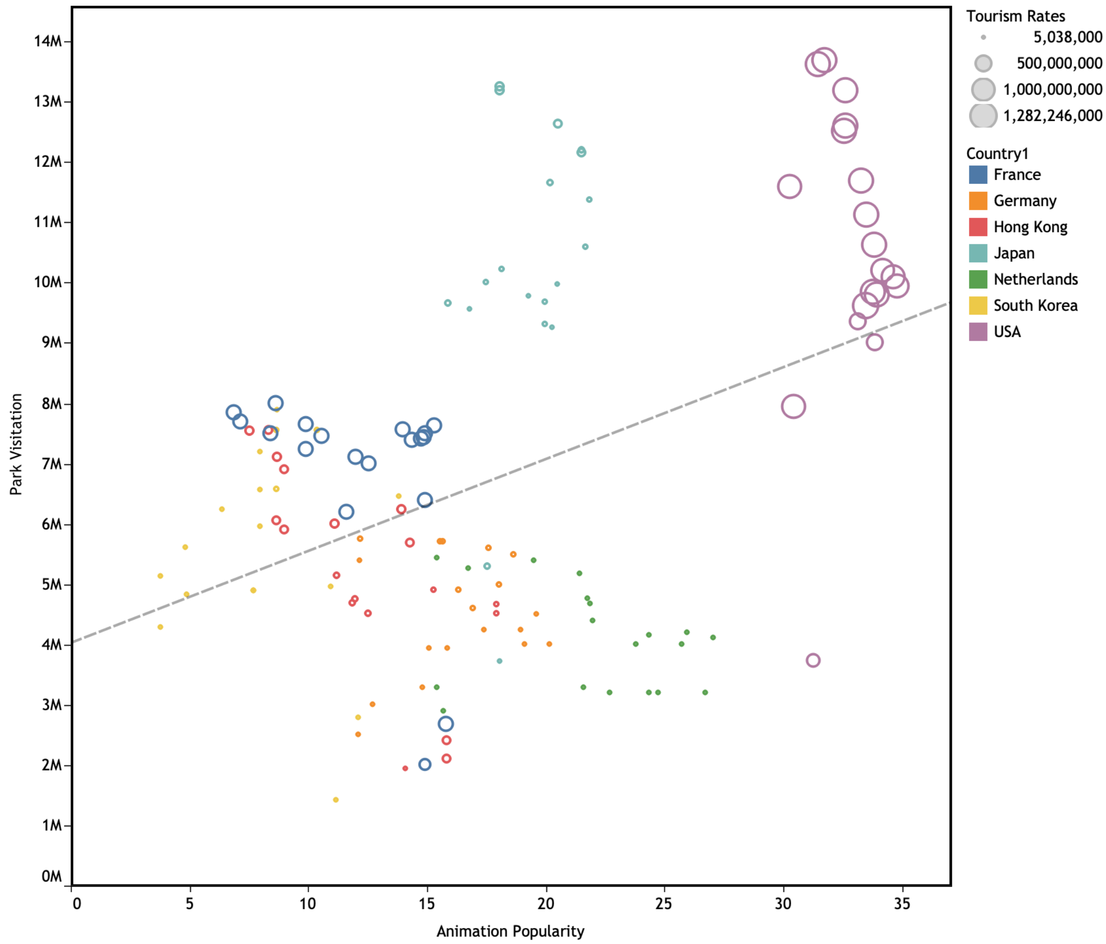


Figure 4.7: Animation Popularity, Visitation, and Tourism)

4.6 Methodology: Panel Data Regression

In this methodology section, I discuss the panel data regression methodology. First, it is essential to mention that the data used in this research is panel data, the dataset was collected manually from several sources including the Park database and The World Bank database. All of this was done in order to analyze how the Variable AnimationPopularity affects theme park visitation rates (Visitation). Moreover, the method of analysis was decided to be a panel data regression, the hypothesis of which implies that the popularity of the animation industry impacts visitation rates by increasing its popularity, which in turn draws tourism to those theme parks and thus stimulates the economy of the hosting country. Additionally, diagnostic checks will be conducted in the subsequent chapter to check the model for issues like Heterskedasticity autocorrelation and multicollinearity, to name a few.

To further proceed with the analysis, as already mentioned, it was determined to use panel data regression for the purposes of establishing the relationship between theme park visitation rates and animation popularity along with several other control variables that reflect economic and demographic factors. Panel data regression was most suitable in this case since the data encompass cross-sectional and time series dimensions. Therefore, the chosen approach enables the control of individual heterogeneity and time-specific effects, which offers a more robust insight into the dynamics among variables over time.

$$\begin{aligned}\textbf{Visitation}_{it} = \beta_0 + \beta_1 &\textbf{AnimationPopularity}_{it} \\ &+ \beta_2 \textbf{Tourism}_{it} + \beta_3 \textbf{GDP}_{it} + \beta_4 \textbf{Population}_{it} + u_{it}\end{aligned} \tag{4.1}$$

Figure 4.8: Panel Data Regression Model

In Figure 4.8, Visitation denotes the dependent variable, representing the rate of visitation at a specific theme park location i during a specific time period t. The independent variables include AnimationPopularity, Tourism, GDP, and Population, which are all indexed by lowercase i and t. The beta coefficients are the parameters to be estimated, which each individually indicates the impact of each and every variable on visitation rates. The error term u at the end of the equation captures any unobserved factors influencing the visitation rates that are not captured by the given model.

Overall, the variables AnimationPopularity, GDP, Population, and Tourism are all independent variables influencing the dependent variable Visitation. AnimationPopularity captures the influence of the popularity of animated content on visitation rates, while tourism reflects the influence of any influx of foreigners in the region and what impact that causes. GDP represents the economic activity of the country, which could potentially influence theme park visitation rates through factors like the availability of disposable income. Population accounts for the size of the country, which could also indicate the potential demand for tourism. The incorporation of these variables into the panel data regression allows for a deeper understanding of the determinants of visitation rates over time and across various geographic locations.

CHAPTER 5

Results

In this chapter, I present my panel data regression results and diagnostic checks.

5.1 Regression Model

Figure 5.1 shows a regression model for panel data with entrees starting from 2004 and going up to 2022. The method used for this regression is Ordinary Least Squares (OLS). Overall, the results of this regression are comprehensive and offer insight into whether theme park visitation rates are influenced by animation popularity. The model has an R-squared of 0.0547, which indicates that around 5.47% of the variability in theme, park visitations can be explained by combining AnimationPopularity, Tourism, GDP, and Population variables. Moreover, the highly statistically significant F-statistic of 6.8826 and p-value of 0.0002 further shed light on the significance of the model.

Further examining the model on an individual predictor basis reveals additional insights. More specifically, the variable AnimationPopularity emerges as a significant predictor with a parameter coefficient equal to 3.111e+05. This suggests that an increase in the popularity of animated content is associated with a substantial increase in theme park visitation rates, which further emphasizes the role of animations in attracting visitors to theme parks.

The variable Tourism also has a positive relationship with theme park visitation rates (Visitation variable). This indicates that a higher number of tourists would be associated with an increase in theme park visitation rates. Moreover, the following statistical evidence can also indicate that a thriving park that hosts a significant number of visitors could potentially be the cause of an increase in tourism rates.

Figure 5.1: PanelOLS Estimation Summary

Variable	**Estimate**	**Std. Error**	**T-statistic**
Dep. Variable: Visitation			
Estimator: PanelOLS			
No. Observations	361		
Date	Sat, Mar 30 2024		
Time	16:00:21		
Cov. Estimator	Unadjusted		
Parameter	**Coefficient**	**Std. Error**	**P-value**
AnimationPopularity	3.111e+05	3.504e+04	0.0000
Tourism	0.0102	0.0037	0.0059
GDP	-3.545e-07	1.247e-07	0.0047
Population	0.0187	0.0091	0.0395
Statistics	**Value**		
R-squared	0.0547		
R-squared (Between)	0.1405		
R-squared (Within)	-0.2106		
Log-likelihood	-6035.1		
F-statistic	6.8826		
P-value	0.0002		
F-statistic (robust)	4.8691		
P-value	0.0025		

On the other hand, Figure 5.1 shows that GDP has a negative relationship with theme park visitation rates. The following is indicated with the parameter coefficient, equal to -3.545e-07. This could potentially mean that, on average, theme park visitation rates tend to decrease as GDP increases. This trend is counterintuitive and thus warrants further investigation to explore any potential variables not included in the regression model that could be causing the GDP variable to be negatively related to the dependent variable.

Additionally, it is worth mentioning that the Population variable has a positive relationship with the dependent variable. Nevertheless, the Population variable has a high p-value in the regression model, indicating low statistical significance. This indicates that while the population might be influencing theme park visitations, there are still other factors not accounted for in this model that are influencing this relationship.

Lastly, R-squared(between) and R-squared (within) have different signs, meaning the model can explain the entity variation. The negative R-squared (between) value might indicate that the model does not effectively capture the variability within individual entities over time. Thus, such an outcome indicates that unobserved factors are present or some time-specific effects that are not adequately accounted for in the current model.

5.2 Diagnostic Checks

Now that the data has been cleaned and organized and the initial model has been assembled to assess the relationship between the dependent variable theme park visitation rates and independent variables AnimationPopularity, Tourism, GDP, and Population, it is important to perform diagnostic checks to make sure that the model is accurately portraying the statistics. Diagnostic checks are a key step in building a reliable model; they play an important role in underlying the linear regression model, enhancing its readability and reliability. The diagnostic checks can bring into light issues like multicollinearity, heteroscedasticity, and outliers. Addressing the mentioned issues early on enhances the robustness of the analysis and ensures that the results from the model are trustworthy.

Figure 5.2: Correlation Matrix

	AnimationPopularity	**Tourism**	**GDP**	**Population**
AnimationPopularity	1.000000	0.496776	0.864348	0.907741
Tourism	0.496776	1.000000	0.599833	0.590237
GDP	0.864348	0.599833	1.000000	0.968551
Population	0.907741	0.590237	0.968551	1.000000

Upon initial diagnostics, multicollinearity was observed between two independent variables. Primarily, it was revealed that GDP and Population have a high level of correlation, which was brought into light by a value of 0.968551 seen in the generated correlation matrix in Figure 5.2. A high correlation suggests a strong linear relationship between the two mentioned independent variables in the model. This finding raises concern about the potential existence of multicollinearity. However, it is also worth noting that it is difficult to differentiate the two variables as they seem important to the research.

Nevertheless, either GDP or Population would have to be potentially excluded from the model. The following decision will ensure that the chosen variables provide unique information and do not present any redundancies, which would allow for a meaningful interpretation of results. Therefore, it was decided to merge the two variables by dividing GDP by Population and thus receiving the variable GDP Per Capita (GDPPerCapita) for future use.

To simplify the diagnostic process, it was determined to start with a simple linear logistic model - PooledOLS. This approach provides a straightforward Ordinary Least Squares model, which will be useful in determining which model to use for the given analysis. However, certain criteria should be met in order to use PooledOLS, such as homoskedasticity, independence, variation over time, etc.

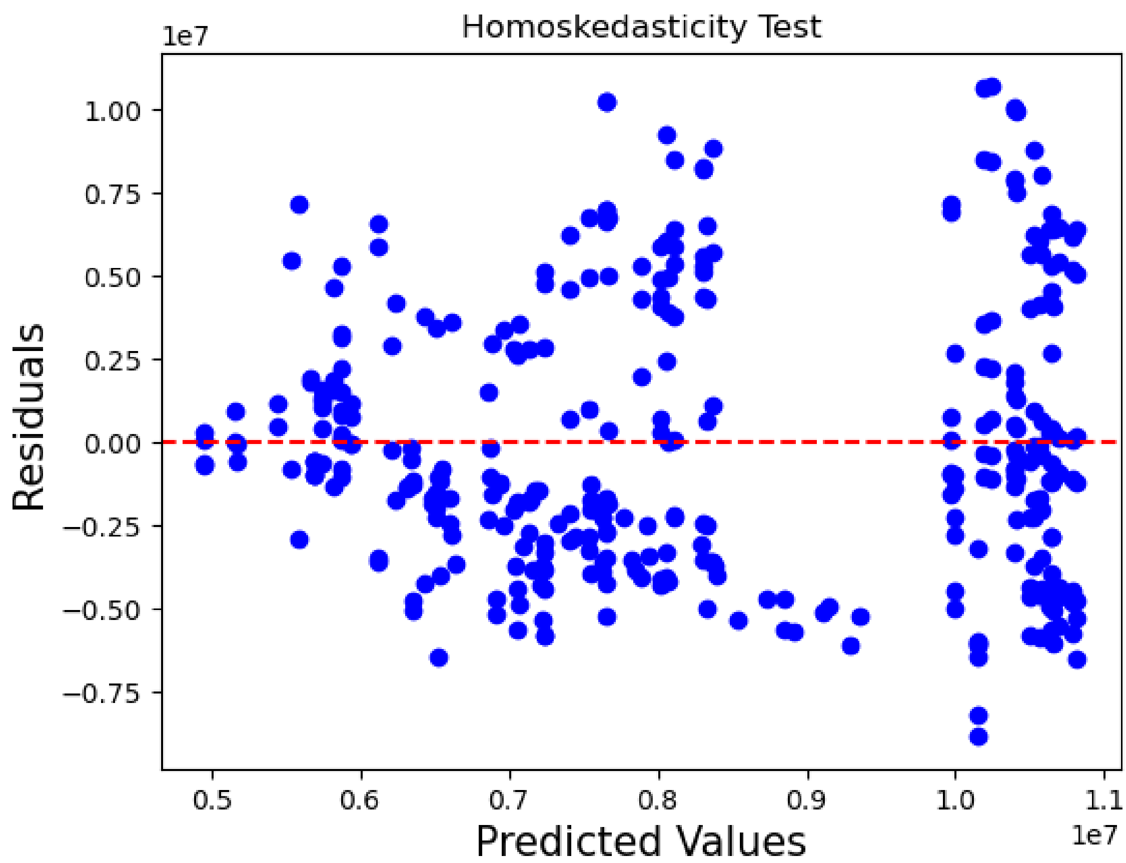


Figure 5.3: Homoskedasticity Test

Moving forward, a homoskedasticity test was performed to ensure the model's robustness, along with ensuring that the model has no autocorrelation. A resides plot created, as seen in Figure 5.3, depicts the predicted value on the x-axis and residuals on the y-axis. The following plot was created to visualize and inspect the spread of data. The observed pattern from the graph indicates a potential violation of homoskedasticity, which suggests growing variance across the predicted values. Therefore, to confirm and quantify this suspicion, both the White and the Breusch Pegan tests were employed. These statistical tests are designed to assess the homoskedasticity of a model. Both the White and Breusch-Pegan test have revealed the values 'LM-Stat': 34.991710772519376, 'LM p-val': 2.5214278739225677e-08, 'F-Stat': 19.21275143991949, 'F p-val': 1.1861453597997972e-08 and 'LM-Stat': 25.34038880235893, 'LM p-val': 4.805407552463173e-07, 'F-Stat': 27.102455215233793, 'F p-val': 3.255283000339051e-07. Both tests have shown significantly low P-values (White test < 1.9 and Breusch-Pegan < 3.3), implying heteroskedasticity. In summary, both of these tests indicate that the condition of homoskedasticity is violated and that there is evidence to support the existence of heteroskedasticity in the residuals. This concludes that the first violation has been proven.

Figure 5.4: Random Effects (RE) Model

Variable	**Estimate**	**Std. Error**	**T-statistic**
Dep. Variable: Visitation			
Estimator: RandomEffects			
No. Observations	361		
Date	Sat, Mar 30 2024		
Time	16:20:36		
Cov. Estimator	Unadjusted		
Parameter	**Coefficient**	**Std. Error**	**P-value**
const	4.973e+06	1.476e+06	0.0008
AnimationPopularity	-6.09e+04	4.307e+04	0.1583
Tourism	0.0410	0.0039	0.0000
Population	0.0070	0.0066	0.2946
Statistics	**Value**		
R-squared	0.2600		
R-squared (Between)	-0.1824		
R-squared (Within)	0.2800		
R-squared (Overall)	-0.0693		
Log-likelihood	-5744.6		
F-statistic	41.808		
P-value	0.0000		
F-statistic (robust)	41.808		
P-value	0.0000		

Figure 5.5: Fixed Effect (FE) Model

Variable	**Estimate**	**Std. Error**	**T-statistic**
Dep. Variable: Visitation			
Estimator: PanelOLS			
No. Observations	361		
Date	Sat, Mar 30 2024		
Time	16:20:37		
Cov. Estimator	Unadjusted		
Parameter	**Coefficient**	**Std. Error**	**P-value**
const	4.591e+06	2.696e+06	1.7029
AnimationPopularity	-6.164e+04	4.749e+04	-1.2979
Tourism	0.0442	0.0039	11.206
Population	0.0077	0.0141	0.5453
Statistics	**Value**		
R-squared	0.2815		
R-squared (Between)	-0.2548		
R-squared (Within)	0.2815		
R-squared (Overall)	-0.1237		
Log-likelihood	-5731.2		
F-statistic	44.280		
P-value	0.0000		

The Durbin-Watson test was employed to test the presence of autocorrelation in the residuals, which produced a statistical value equal to 0.41. The Durbin-Watson test statistics range from 0 to 4, where a value of 2 indicates no autocorrelation. Values between 0 and 2 imply positive autocorrelation, and the lower values indicate strong autocorrelation. At the same time, values between 2 and 4 are negative autocorrelations, with higher values indicating more substantial autocorrelation. In the given analysis, the Durbin-Watson statistics is equal to 0.41, which means a strong negative autocorrelation. The outputted small value (<2) indicates a substantial number of residuals that go unaccounted for by the model. Given these findings, Fixed Effect (FE) and Random Effect (RE) models would be more applicable to address the identified autocorrelation and heteroskedasticity and enhance the overall robustness of the model.

As seen in Figure 5.4, the Random Effects (RE) model provides insight into the panel data regression model, indicating a reported R-squared of 0.26, indicating that roughly 26% of observations are being predicted. The analysis suggests that AnimationPopularity does not necessarily substantially affect Visitation rates, whereas Tourism exhibits a strong positive relationship. On the other hand, Population does not initially appear to be significant. In contrast, the fixed effect (FE) model seen in Figure 5.5 showcased an improved R-squared of 0.2815, which suggests that approximately 28% of theme park visitations are predicted by the model; the F-test score indicates the presence of entity effects. The parameter estimates in both models show similarity in AnimationPopularity and Tourism while Population remains not very significant. Including entity-specific effects in the PanelOLS model accounts for unobserved heterogeneity, potentially increasing the model's explanatory ability.

When comparing the two models even closely, it can be observed that they differ in their treatment of entity-specific effects. The Random Effects model overlooks entity-specific heterogeneity, which is evident with an R-squared value of 0.2815, meaning that the results might be the production of biased estimates. On the contrary, the PanelOLS model accounts, which has a higher R-squared value of 0.2815, account for entity-specific effects, which leads to a more comprehensive understanding of the relationship between the independent and dependent variables. The F-test in the PanelOLS supports the inclusion of entity-specific effects, emphasizing the importance of accounting for unobserved heterogeneity. In summary, both models provide valuable insight; nevertheless, the PanelOLS model adds an advantage by considering entity-specific effects, so the FE model proves to be more suitable since endogeneity is present in the model.

5.3 Adjusting the Model After the Diagnostic Checks

After conducting the diagnostic checks mentioned in the previous section, it was determined that the variables GDP and Populations should be permanently removed due to multicollinearity-related issues. The two were highly correlated, so instead of GDP and Population, a new variable was introduced called GDPPerCapita, which represents the GDP variable divided by the population variable. Additionally, the previous diagnostic check revealed quite a high constant associated with the variable Visitation; therefore, it was decided to log this variable as the solution to this issue.

After the adjustments mentioned above were made, the Ordinary Least Squares regression model generated seen in Figure 5.6 showed an R squared value equaling 0.157, which indicated that the model approximately predicted 15.7% of outcomes in the logged visitation rates, which means that the variation by the independent variables could have somewhat of a predicting power. Moreover, the adjusted are squared was observed to be equal to 0.150.

Figure 5.6: Regression Mode With log(Visitation)

Variable	**Coefficient**	**Std. Error**	**T-statistic**
Dep. Variable: Log_Visitation			
Model: OLS			
No. Observations	361		
Date	Sat, 30 Mar 2024		
Time	16:36:30		
Parameter	**Value**	**Std. Error**	**P-value**
const	15.3715	0.124	0.000
AnimationPopularity	0.0225	0.005	0.000
Tourism	1.304e-09	4.66e-10	0.005
GDPPerCapita	-4.447e-06	3.78e-06	0.240
Statistics	**Value**		
R-squared	0.157		
Adj. R-squared	0.150		
F-statistic	22.17		
Prob (F-statistic)	3.45e-13		
Log-Likelihood	-307.91		
AIC	623.8		
BIC	639.4		
Omnibus	136.013		
Durbin-Watson	0.691		
Jarque-Bera (JB)	1006.910		
Prob(JB)	2.25e-219		

The new model's F statistic is 22.17, and its probability value is 3.45e-13. These two results indicate that the overall regression model is significant. Therefore, the overall model is statistically significant. This suggests that at least one variable in the model among the independent variables has a non-zero coefficient, which explains the variation in the dependent variable.

Additionally, after examining the coefficients associated with each variable, it was concluded that AnimationPopularity and Tourism have p-values well below 0.05, indicating statistical significance. On the other hand, the coefficient of GDPPerCapita was found to be statistically insignificant, with a p-value of 0.24. This indicates that GDP per capita does not affect the logged version of visitation rates.

Figure 5.7: Variance Inflation Factor (VIF), Homoskedasticity Test, and Normality Test

Feature	**VIF**
const	17.0619
AnimationPopularity	2.0966
Tourism	1.4191
GDPPerCapita	2.1239
Homoskedasticity Test (Breusch-Pagan)	
p-value	0.4806
Normality Test (Jarque-Bera)	
p-value	2.92×10^{-30}

Furthermore, diagnostic checks have been performed again to check the strength and validity of the new model, as seen in Figure 5.7. These tests included a multicollinearity test, a heteroskedasticity test, and the normality test. Therefore, the results of these tests indicated some evidence that the model suffers from heteroskedasticity and that the data is not normally distributed. With that being said, the new model does indeed show progress compared to the previous one by fixing the issue with multicollinearity; however, to develop the model further and take a step towards better results, it was determined to log GDPPerCapita and Tourism since their coefficients were higher compared to AnimationPopularity.

Figure 5.8: New Modified Regression Model

Variable	**Estimate**	**Std. Error**	**T-statistic**
Dep. Variable: Visitation			
Estimator: OLS			
No. Observations	361		
R-squared	0.157		
Adj. R-squared	0.150		
F-statistic	22.17		
P-value (F-stat)	3.45e-13		
Parameter	**Coefficient**	**Std. Error**	**P-value**
const	15.3715	0.124	0.000
AnimationPopularity	0.0225	0.005	0.000
Tourism	1.304e-09	4.66e-10	0.005
GDPPerCapita	-4.447e-06	3.78e-06	0.240
VIF			
const	17.061931		
AnimationPopularity	2.096623		
Tourism	1.419107		
GDPPerCapita	2.123934		
Homoskedasticity Test			
p-value	0.480620		
Normality Test			
p-value	2.92×10^{-30}		

After the second stage, diagnostic checks were performed, and additional modifications were made to the model to enhance its robustness. More specifically, during the proviso diagnostic checks, it was determined that Tourism and GDpPerCapita should not be logged along with Visitation, which was logged before.

AnimationPopulati was not logged since the coefficients were small enough that they were already a good fit for the model.

After examining the regression results seen in Figure 5.8, it can be observed that the R squared value remained somewhat unchanged at 0.156, which indicates that in the new model, approximately 15.6% of the variation in the logged version of Visitation is explained by the independent variables included in the model. The adjusted R squared value also remained relatively unchanged and is equal to 0.149 in the new model.

Additionally, the F statistic in the new model is equal to 21.96, and the corresponding probability value is 4.49e-13, meaning that the model remained statistically significant, which again indicates that at least one of the independent variables has a nonzero coefficient, which explains the variation in the dependent variable. However, it is also worth mentioning that the F statistic has decreased slightly compared to the previous model, indicating marginally less statistical significance. Further examination of the coefficients of individual variables revealed that the log version of tourism and AnimationPopularity remained statistically significant at a 5% level, while the log version of GDP per capita is still statistically insignificant with a p-value of 0.193. This implies that the log version of GDPPerCapita does not have any predictive power over the logged version of visitation.

To further prove the robustness of the model, diagnostic checks were performed once again, the results of which can be seen in Figure 5.9. The diagnostic checks showed that the model is exhibiting some issues. Still, the overall conclusion can be that it has increased its validity compared to the previous model. This is evident in its lack of exhibiting any issues with multicollinearity and heteroskedasticity.

Overall, while the adjustments and modifications have improved the model, some slight modifications still need to be performed for further improvement. Mainly, it is essential to find a number of additional independent variables that would increase the model's value by accounting for a greater variance in theme park visitation rates. Nevertheless, the overall result can still be concluded to be significant enough to prove the relationship between the popularity of animated content and theme park visitation rates.

Figure 5.9: New Variance Inflation Factor (VIF), Homoskedasticity Test, and Normality Test

Feature	**VIF**
const	2848.893890
AnimationPopularity	2.127741
Log_Tourism	1.517277
Log_GDPPerCapita	2.227900
Homoskedasticity Test (Breusch-Pagan)	
p-value	0.251303
Normality Test (Jarque-Bera)	
p-value	1.50×10^{-29}

CHAPTER 6

Discussion and Conclusion

The study explored the relationship between animation popularity and theme park visitation rates. It attempts to prove that the animation industry can be used to market different regions, attract tourists to those regions, and, later on, increase economic activity. The dataset used for analyzing this relationship encompassed 19 theme parks from all around the globe. The main variables identified were AnimationPopularity, Park, and Visitations. In contrast, the control variables chosen were GDP, Population, and Tourism, which helped to account for any impact demographic or economic factors might have had on the matter.

The data was initially cleaning addressed the 51 missing values in the variable Visitation and Tourism, ensuring the integrity and robustness of the dataset. Additionally, the explanatory data analysis that followed the data cleaning uncovered trends within the global animation and entertainment industries. Mainly, the selected theme parks worldwide exhibited an increasing trend from 2004 to 2020, followed by a sharp downturn due to the coronavirus pandemic, after which the visitation trend started to increase again. Overall, the trend revealed the industry's strong growth and resilience over 18 years.

Moving forward, analyzing theme park visitation trends on a country-by-country basis sheds light on some regional disparities. The USA, Japan, and France emerged with the most theme park visits. On the other hand, a parallel examination of global tourism rates revealed that the USA, France, and Hong Kong are the leading countries hosting the most tourists annually. The main factor leading this analysis was the relationship between GDP, tourism rates, animation popularity, and theme park visitation rates. The exploratory data analysis revealed a statistically significant relationship, shedding light on the importance of animation popularity in drawing visitors to theme parks. The positive correlation between animation popularity and visitation rates, along with the notable impact of tourism rates, justified the notion that theme parks featuring well-known cartoon characters have the potential to attract overseas visitors.

The regression model employed in this research paper utilized Ordinary Least Squares (OLS), which had mixed findings compared to the exploratory data analysis findings. AnimationPopularity emerged as a significant predictor helping to influence and predict theme park visitation rates. However, it is also worth mentioning that the inclusion of a control variable included a counterintuitive negative relationship between GDP and theme park visitation rates, which calls for further investigation to help recognize potentially unaccounted factors.

Moreover, diagnostics checks have brought into light some challenges, such as heteroskedasticity, multicollinearity, and autocorrelation, which proved the initial model not to be as effective and prompted to switch to the Random Effects (RE) and Fixed Effects (FE) models. On top of that, while both models offered valuable insight, the PanleOLS mode proved to be more helpful and suitable due to the fact that it accounts for entity effects, which addressed the issue with autocorrelation and heteroskedasticity.

In conclusion, the study attempts to understand the nuanced relationship theme park visitation rates have with animation popularity to try to answer the question of whether animation popularity increases economic activity. Moreover, from the analysis, the interplay between theme park visitations, animation popularity, and tourism rates further underscores the complex nature of the animation industry. As a recommendation for future research on the given topic, it is essential to delve deeper into the relationship between economic activity and animation popularity by creating a more complex data source that would include a more extensive array of possible control variables to help create better and more reliable analysis results.

APPENDIX A

Python Code Utilized

Listing A.1: Packages Used

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.impute import SimpleImputer
import numpy as np
from linearmodels.panel import PanelOLS
```

```
df = pd.read_excel('/Users/davidsokurov/Desktop/I.S./Data Analysis/Final I.S. Dataset.xlsx')
```

Listing A.2: Initial Data Cleaning Process

```
# Total missing values are 51 (Visitation and Tourism)
missing_values = df.isnull().sum()
print(missing_values)
total_missing_values = df.isnull().sum().sum()

print("Total missing values:", total_missing_values)
```

Output

```
Park                    0
Country                 0
Year                    0
Visitation             13
AnimatioPopularity      0
GDP                     0
Population              0
Tourism                38
dtype: int64
Total missing values: 51
```

```python
# Get the list of variables
variable_list = df.columns.tolist()
print("List of variables:", variable_list)
```

```python
unique_parks = df['Park'].unique()
unique_park_count = df['Park'].nunique()
print(unique_park_count)
print(unique_parks)
```

```python
# Gives the list of all missing values and their locations
missing_values_df = df.isnull()

missing_locations = pd.DataFrame([(row, col) for row in
    missing_values_df.index for col in missing_values_df.
    columns if missing_values_df.at[row, col]])
print("Locations of missing values:")
print(missing_locations)
```

```
# Imputes the variables Visitation and Tourism
df['Visitation'] = df.groupby('Park')['Visitation'].
    transform(lambda x: x.fillna(x.mean()))
df['Tourism'] = df.groupby('Park')['Tourism'].transform(
    lambda x: x.fillna(x.mean()))
```

```
#df.to_excel("edited_dataset.xlsx")
unique_countries = df["Country"].unique()
print(unique_countries)
```

Listing A.3: Exploratory Data Anlaysis

```
plt.figure(figsize=(15, 6))
for park, data in df.groupby('Park'):
    plt.plot(data['Year'], data['Visitation'], label=park)
plt.title('Visitation Trends Over the Years')
plt.xlabel('Year')
plt.ylabel('Visitation')
plt.legend(loc='upper left')
plt.show()
```

```
mean_visitation_by_country = df.groupby('Country')['
    Visitation'].mean()

# Sort countries in ascending order based on mean
    visitation
sorted_countries = mean_visitation_by_country.sort_values
    ().index
```

```

plt.figure(figsize=(20, 10))
for country in sorted_countries:
    plt.bar(country, mean_visitation_by_country[country],
        label=country)
    plt.text(country, mean_visitation_by_country[country],
         f'{mean_visitation_by_country[country]:.2f}', ha='
        center', va='bottom')

plt.title('Visitation Across Countries (Ascending Order)')
plt.xlabel('Country')
plt.ylabel('Mean Visitation')
plt.legend()
plt.show()
```

```
plt.figure(figsize=(15, 8))
df.groupby('Year')['Visitation'].mean().plot(marker='o')
plt.title('Mean Visitation Over the Years')
plt.xlabel('Year')
plt.ylabel('Mean Visitation')
plt.show()
```

```
plt.figure(figsize=(15, 8))
tourism_by_country = df.groupby('Country')['Tourism'].sum
    ().sort_values(ascending=True)
tourism_by_country.plot(kind='barh', color='skyblue')

# Add actual values on top of the bars
for index, value in enumerate(tourism_by_country):
```

```
    plt.text(value, index, f'{value:,}', va='center',
        fontsize=10)

plt.title('Total Tourism by Country')
plt.xlabel('Total Tourism')
plt.ylabel('Country')
plt.show()
```

```
# Get the list of unique countries
countries = df['Country'].unique()

# Create subplots for each country
fig, axes = plt.subplots(nrows=len(countries), ncols=1,
    figsize=(10, 3 * len(countries)))

# Iterate over countries and create scatter plots with
    trend lines
for i, country in enumerate(countries):
    country_data = df[df['Country'] == country]

    # Scatter Plot: Animation Popularity vs. Visitation
    axes[i].scatter(x='AnimationPopularity', y='Visitation
        ', data=country_data, alpha=0.5, label='Visitation'
        )

    # Trend line for Animation Popularity vs. Visitation
    z_visitation = np.polyfit(country_data['
        AnimationPopularity'], country_data['Visitation'],
        1)
```

```
    p_visitation = np.poly1d(z_visitation)
    axes[i].plot(country_data['AnimationPopularity'], p_visitation(country_data['AnimationPopularity']), color='red')

    # Scatter Plot: Animation Popularity vs. Tourism
    axes[i].scatter(x='AnimationPopularity', y='Tourism', data=country_data, alpha=0.5, label='Tourism')

    # Trend line for Animation Popularity vs. Tourism
    z_tourism = np.polyfit(country_data['AnimationPopularity'], country_data['Tourism'], 1)
    p_tourism = np.poly1d(z_tourism)
    axes[i].plot(country_data['AnimationPopularity'], p_tourism(country_data['AnimationPopularity']), color='blue')

    axes[i].set_title(f'Scatter Plot: Animation Popularity, Visitation, and Tourism - {country}')
    axes[i].set_xlabel('Animation Popularity')
    axes[i].set_ylabel('Count')
    axes[i].legend()

plt.tight_layout()
plt.show()
```

Listing A.4: Initial Regerssion Model

```
df['Year'] = pd.to_datetime(df['Year'], format='%Y')
```

```python

# Set the DataFrame index using the panel variables and sort it
df.set_index(['Park', 'Year'], inplace=True)
df.sort_index(inplace=True)

# Create a PanelOLS model
model = PanelOLS.from_formula('Visitation ~ AnimationPopularity + Tourism + GDP + Population', data=df)

# Fit the model
result = model.fit()

# Print regression results
print(result)
```

Output

```
# Output of the code block
 PanelOLS Estimation Summary
================================================================================
Dep. Variable:           Visitation  R-squared:                        0.0547
Estimator:                 PanelOLS  R-squared (Between):              0.1405
No. Observations:               361  R-squared (Within):              -0.2106
Date:              Sun, Mar 10 2024  R-squared (Overall):              0.0547
Time:                      18:29:17  Log-likelihood                   -6035.1
Cov. Estimator:          Unadjusted
                                      F-statistic:                      6.8826
Entities:                        19  P-value                           0.0002
Avg Obs:                     19.000  Distribution:                   F(3,357)
Min Obs:                     19.000
Max Obs:                     19.000  F-statistic (robust):             4.8691
                                      P-value                           0.0025
Time periods:                    19  Distribution:                   F(3,357)
Avg Obs:                     19.000
Min Obs:                     19.000
Max Obs:                     19.000
```

Output

```
Parameter Estimates
=======================================================
                   Parameter  Std. Err.   T-stat P-value
-------------------------------------------------------
AnimationPopularity 3.111e+05 3.504e+04   8.8793 0.0000
Tourism              0.0102    0.0037     2.7699 0.0059
GDP                -3.545e-07 1.247e-07 -2.8418 0.0047
Population           0.0187    0.0091     2.0663 0.0395
=======================================================
```

Listing A.5: Diagnostic Checks and Updated Regression Model

```python
df = pd.read_excel('/Users/davidsokurov/Desktop/I.S./Data Analysis/Final I.S. Dataset.xlsx')
df.set_index(["Park", "Year"], inplace=True)



df['Visitation'] = df.groupby('Park')['Visitation'].transform(lambda x: x.fillna(x.mean()))
```

```
df['Tourism'] = df.groupby('Park')['Tourism'].transform(
    lambda x: x.fillna(x.mean()))
```

```
years = df.index.get_level_values("Year").to_list()
df["Year"] = pd.Categorical(years)
```

```
correlation_matrix = df[["AnimationPopularity", "Tourism",
     "GDP", "Population"]].corr()
print(correlation_matrix)
```

Output

```
                     AnimationPopularity  Tourism    GDP  Population
AnimationPopularity             1.000000  0.496776  0.864348  0.907741
Tourism                         0.496776  1.000000  0.599833  0.590237
GDP                             0.864348  0.599833  1.000000  0.968551
Population                      0.907741  0.590237  0.968551  1.000000
```

```
# Perform PooledOLS
from linearmodels import PooledOLS
import statsmodels.api as sm
exog = sm.tools.tools.add_constant(df["AnimationPopularity
    "])

endog = df["Visitation"]
mod = PooledOLS(endog, exog)
pooledOLS_res = mod.fit(cov_type='clustered',
    cluster_entity=True)
# Store values for checking homoskedasticity graphically
fittedvals_pooled_OLS = pooledOLS_res.predict().
    fitted_values
residuals_pooled_OLS = pooledOLS_res.resids
```

```
# 3A. Homoskedasticity
import matplotlib.pyplot as plt
 # 3A.1 Residuals-Plot for growing Variance Detection
fig, ax = plt.subplots()
ax.scatter(fittedvals_pooled_OLS, residuals_pooled_OLS,
    color = "blue")
ax.axhline(0, color = 'r', ls = '--')
ax.set_xlabel("Predicted Values", fontsize = 15)
ax.set_ylabel("Residuals", fontsize = 15)
ax.set_title("Homoskedasticity Test", fontsize = 30)
plt.show()
```

```
# 3A.2 White-Test
from statsmodels.stats.diagnostic import het_white,
    het_breuschpagan

# Assuming you have residuals_pooled_OLS from your
    previous code
pooled_OLS_dataset = pd.concat([df, residuals_pooled_OLS.
    rename("residuals")], axis=1)
pooled_OLS_dataset = pooled_OLS_dataset.drop(["Year"],
    axis=1).fillna(0)
exog = sm.tools.tools.add_constant(df["AnimationPopularity
    "]).fillna(0)
white_test_results = het_white(pooled_OLS_dataset["
    residuals"], exog)
labels = ["LM-Stat", "LM p-val", "F-Stat", "F p-val"]
print(dict(zip(labels, white_test_results)))
```

```

# 3A.3 Breusch-Pagan-Test
breusch_pagan_test_results = het_breuschpagan(
    pooled_OLS_dataset["residuals"], exog)
labels = ["LM-Stat", "LM p-val", "F-Stat", "F p-val"]
print(dict(zip(labels, breusch_pagan_test_results)))
```

Output

```
{'LM-Stat': 34.991710772519376,
'LM p-val': 2.5214278739225677e-08,
'F-Stat': 19.21275143991949, 'F p-val': 1.1861453597997972e-08}
{'LM-Stat': 25.34038880235893,
'LM p-val': 4.805407552463173e-07,
'F-Stat': 27.102455215233793, 'F p-val': 3.255283000339051e-07}
```

```
# 3.B Non-Autocorrelation
# Durbin-Watson-Test
from statsmodels.stats.stattools import durbin_watson

durbin_watson_test_results = durbin_watson(
    pooled_OLS_dataset["residuals"])
print(durbin_watson_test_results)
```

```
from linearmodels import PanelOLS
from linearmodels import RandomEffects
import statsmodels.api as sm

```

```
# Assuming df contains 'AnimationPopularity', 'Tourism', 'GDP', 'Population', and 'Visitation'
exog_vars = ['AnimationPopularity', 'Tourism', 'Population']
exog = sm.tools.tools.add_constant(df[exog_vars])
endog = df['Visitation']

# Random effects model
model_re = RandomEffects(endog, exog, check_rank=False)
re_res = model_re.fit()

# Fixed effects model
model_fe = PanelOLS(endog, exog, entity_effects=True)
fe_res = model_fe.fit()

# Print results
print(re_res)
```

RandomEffects Estimation Summary

```
================================================================================
Dep. Variable:             Visitation   R-squared:                        0.2600
Estimator:              RandomEffects   R-squared (Between):             -0.1824
No. Observations:                 361   R-squared (Within):               0.2800
Date:                Sat, Mar 30 2024   R-squared (Overall):             -0.0693
Time:                        16:20:36   Log-likelihood                   -5744.6
Cov. Estimator:            Unadjusted
                                        F-statistic:                      41.808
Entities:                          19   P-value                           0.0000
Avg Obs:                       19.000   Distribution:                   F(3,357)
Min Obs:                       19.000
Max Obs:                       19.000   F-statistic (robust):             41.808
                                        P-value                           0.0000
Time periods:                      19   Distribution:                   F(3,357)
Avg Obs:                       19.000
Min Obs:                       19.000
Max Obs:                       19.000

                                  Parameter Estimates
=======================================================================================
                     Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
---------------------------------------------------------------------------------------
const                4.973e+06  1.476e+06     3.3696     0.0008    2.07e+06   7.875e+06
AnimationPopularity  -6.09e+04  4.307e+04    -1.4138     0.1583  -1.456e+05   2.381e+04
Tourism                 0.0410     0.0039     10.617     0.0000      0.0334      0.0486
Population              0.0070     0.0066     1.0496     0.2946     -0.0061      0.0200
=======================================================================================
```

```
print(fe_res)
```

PanelOLS Estimation Summary

```
================================================================================
Dep. Variable:             Visitation   R-squared:                        0.2815
Estimator:                   PanelOLS   R-squared (Between):             -0.2548
No. Observations:                 361   R-squared (Within):               0.2815
Date:                Sat, Mar 30 2024   R-squared (Overall):             -0.1237
Time:                        16:20:37   Log-likelihood                   -5731.2
Cov. Estimator:            Unadjusted
                                        F-statistic:                      44.280
Entities:                          19   P-value                           0.0000
Avg Obs:                       19.000   Distribution:                   F(3,339)
Min Obs:                       19.000
Max Obs:                       19.000   F-statistic (robust):             44.280
                                        P-value                           0.0000
Time periods:                      19   Distribution:                   F(3,339)
Avg Obs:                       19.000
Min Obs:                       19.000
Max Obs:                       19.000

                                  Parameter Estimates
=======================================================================================
                     Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
---------------------------------------------------------------------------------------
const                4.591e+06  2.696e+06     1.7029     0.0895  -7.122e+05   9.895e+06
AnimationPopularity -6.164e+04  4.749e+04    -1.2979     0.1952  -1.551e+05   3.178e+04
Tourism                 0.0442     0.0039     11.206     0.0000      0.0365      0.0520
Population              0.0077     0.0141     0.5453     0.5859     -0.0200      0.0353
=======================================================================================

F-test for Poolability: 64.786
P-value: 0.0000
Distribution: F(18,339)

Included effects: Entity
```

```
import pandas as pd
import numpy as np
import statsmodels.api as sm
```

```
from statsmodels.stats.outliers_influence import
    variance_inflation_factor
from statsmodels.compat import lzip
from statsmodels.stats.diagnostic import het_breuschpagan,
    normal_ad
from scipy import stats

# Log transformation of the dependent variable
df['Log_Visitation'] = np.log(df['Visitation'])

# Define independent and dependent variables
X = df[['AnimationPopularity', 'Tourism', 'GDPPerCapita'
    ,]]
y = df['Log_Visitation']

# Add a constant term to the independent variables
X = sm.add_constant(X)

# Fit the regression model
model = sm.OLS(y, X).fit()

# Print the summary of the regression results
print(model.summary())

# Test for multicollinearity
def calculate_vif(X):
    vif_data = pd.DataFrame()
    vif_data['Feature'] = X.columns
```

```
    vif_data['VIF'] = [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]
    return vif_data

print("\nVariance Inflation Factor (VIF):")
print(calculate_vif(X))

# Test for homoskedasticity
_, p_homoskedasticity, _, _ = het_breuschpagan(model.resid, X)
print("\nHomoskedasticity Test (Breusch-Pagan):")
print("p-value:", p_homoskedasticity)

# Test for normality of residuals
p_normality = stats.normaltest(model.resid)[1]
print("\nNormality Test (Jarque-Bera):")
print("p-value:", p_normality)
```

OLS Regression Results

```
==============================================================================
Dep. Variable:          Log_Visitation   R-squared:                       0.157
Model:                             OLS   Adj. R-squared:                  0.150
Method:                  Least Squares   F-statistic:                     22.17
Date:                 Sat, 30 Mar 2024   Prob (F-statistic):           3.45e-13
Time:                         16:36:30   Log-Likelihood:                -307.91
No. Observations:                  361   AIC:                             623.8
Df Residuals:                      357   BIC:                             639.4
Df Model:                            3
Covariance Type:             nonrobust
=======================================================================================
                          coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------
const                  15.3715      0.124    123.837      0.000      15.127      15.616
AnimationPopularity     0.0225      0.005      4.971      0.000       0.014       0.031
Tourism              1.304e-09   4.66e-10      2.799      0.005    3.88e-10    2.22e-09
GDPPerCapita        -4.447e-06   3.78e-06     -1.176      0.240   -1.19e-05    2.99e-06
==============================================================================
Omnibus:                       136.013   Durbin-Watson:                   0.691
Prob(Omnibus):                   0.000   Jarque-Bera (JB):             1006.910
Skew:                           -1.382   Prob(JB):                    2.25e-219
Kurtosis:                       10.701   Cond. No.                     4.79e+08
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.79e+08. This might indicate that there are
strong multicollinearity or other numerical problems.

Variance Inflation Factor (VIF):
               Feature        VIF
0                const  17.061931
1  AnimationPopularity   2.096623
2              Tourism   1.419107
3         GDPPerCapita   2.123934

Homoskedasticity Test (Breusch-Pagan):
p-value: 0.48062047267207664

Normality Test (Jarque-Bera):
p-value: 2.918925958447715e-30
```

```
import pandas as pd
import numpy as np
import statsmodels.api as sm
from statsmodels.stats.outliers_influence import
    variance_inflation_factor
from statsmodels.compat import lzip
from statsmodels.stats.diagnostic import het_breuschpagan,
     normal_ad
from scipy import stats

# Log transformation of the dependent variable
df['Log_Visitation'] = np.log(df['Visitation'])

# Log transformation of independent variables
df['Log_Tourism'] = np.log(df['Tourism'])
df['Log_GDPPerCapita'] = np.log(df['GDPPerCapita'])

# Define independent and dependent variables
X = df[['AnimationPopularity', 'Log_Tourism', '
    Log_GDPPerCapita']]
y = df['Log_Visitation']

# Add a constant term to the independent variables
X = sm.add_constant(X)

# Fit the regression model
model = sm.OLS(y, X).fit()

```

```
# Print the summary of the regression results
print(model.summary())

# Test for multicollinearity
def calculate_vif(X):
    vif_data = pd.DataFrame()
    vif_data['Feature'] = X.columns
    vif_data['VIF'] = [variance_inflation_factor(X.values,
        i) for i in range(X.shape[1])]
    return vif_data

print("\nVariance␣Inflation␣Factor␣(VIF):")
print(calculate_vif(X))

# Test for homoskedasticity
_, p_homoskedasticity, _, _ = het_breuschpagan(model.resid
    , X)
print("\nHomoskedasticity␣Test␣(Breusch-Pagan):")
print("p-value:", p_homoskedasticity)

# Test for normality of residuals
p_normality = stats.normaltest(model.resid)[1]
print("\nNormality␣Test␣(Jarque-Bera):")
print("p-value:", p_normality)
```

OLS Regression Results

```
                            OLS Regression Results
==============================================================================
Dep. Variable:        Log_Visitation   R-squared:                       0.156
Model:                           OLS   Adj. R-squared:                  0.149
Method:                Least Squares   F-statistic:                     21.96
Date:               Sat, 30 Mar 2024   Prob (F-statistic):           4.49e-13
Time:                       17:24:21   Log-Likelihood:                -308.18
No. Observations:                361   AIC:                             624.4
Df Residuals:                    357   BIC:                             639.9
Df Model:                          3
Covariance Type:           nonrobust
```

OLS Regression Results

```
                          coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------
const                  16.2131      1.605     10.101      0.000      13.056      19.370
AnimationPopularity     0.0228      0.005      4.995      0.000       0.014       0.032
Log_Tourism             0.0767      0.028      2.693      0.007       0.021       0.133
Log_GDPPerCapita       -0.2141      0.164     -1.303      0.193      -0.537       0.109
==========================================================================================
Omnibus:                     132.740   Durbin-Watson:                   0.687
Prob(Omnibus):                 0.000   Jarque-Bera (JB):              968.885
Skew:                         -1.346   Prob(JB):                    4.07e-211
Kurtosis:                     10.561   Cond. No.                     1.65e+03
==========================================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified
[2] The condition number is large, 1.65e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Output

```
Variance Inflation Factor (VIF):
                Feature          VIF
0                 const  2848.893890
1  AnimationPopularity     2.127741
2           Log_Tourism     1.517277
3       Log_GDPPerCapita     2.227900


Homoskedasticity Test (Breusch-Pagan):
p-value: 0.25130333442270303


Normality Test (Jarque-Bera):
p-value: 1.4993670420555388e-29
```

References

[1] N. Abd Rahman, Z.A. Dawam, and J.K. Chan. "The Characteristics of Film Products to Induce Tourism". In: *Journal of Tourism, Hospitality and Environment Management* 4.16 (2019), pp. 84–99.

[2] C. Baker. "The Prevalence of Storyworlds and Thematic Landscapes in Global Theme Parks". In: *Annals of Tourism Research Empirical Insights* 4.1 (2023), p. 100080.

[3] S. Beeton and P. Seaton. "Creating Places and Transferring Culture: American Theme Parks in Japan". In: *Film Tourism in Asia. Perspectives on Asian Tourism*. Ed. by S. Kim and S. Reijnders. 2018, Springer, Singapore.

[4] T. Brown. "The DVD of Attractions? The Lion King and the Digital Theme Park". In: *Convergence: The International Journal of Research into New Media Technologies* 13.2 (2007), pp. 169–183.

[5] N.T. Bu et al. "Creating a Re-Tech Theme Park Through Neuromarketing and Cartoon Characters". In: (2022). URL: https://www.preprints.org/manuscript/202203.0677/v1.

[6] D. Buhalis et al. "Technological Disruptions in Services: Lessons from Tourism and Hospitality". In: *Journal of Service Management* 30.4 (2019), pp. 484–506.

[7] Q.P. Chen, J.J. Wu, and W.Q. Ruan. "What Fascinates You? Structural Dimension and Element Analysis of Sensory Impressions of Tourist Destinations

Created by Animated Works". In: *Asia Pacific Journal of Tourism Research* 26.9 (2021), pp. 1038–1054.

[8] F. Florian. "Like Walking Into a Movie: Intermedial Relations Between Theme Parks and Movies". In: *The Journal of Popular Culture* 50.4 (2017), pp. 704–722.

[9] L. Florido-Benítez. "International Mobile Marketing: A Satisfactory Concept for Companies and Users in Times of Pandemic". In: *Benchmarking: An International Journal* 29.6 (2022), pp. 1826–1856.

[10] Lázaro Florido-Benítez. "Film-induced Tourism – The Impact of Animation, Cartoon, Superhero, and Fantasy Movies". In: *Tourism Review* 78.5 (2023), pp. 1298–1314. URL: https://doi.org/10.1108/tr-11-2022-0537.

[11] Hiroaki Mori. "The Power of Anime: A New Driver of Volunteer Tourism". In: *Tourism and Hospitality* 3.2 (2022), pp. 330–344. URL: https://doi.org/10.3390/tourhosp3020022.

[12] Platform, Other, and Joint Secretariat. *The Booming Market of the Animation Industry*. World Wide Web. 2023. URL: https://www.interregeurope.eu/find-policy-solutions/stories/the-booming-market-of-the-animation-industry.

[13] W. Pompl. "The Concept of Animation: Aspects of Tourism Services". In: *Tourism Management* 4.1 (1983), pp. 3–11. URL: https://doi.org/10.1016/0261-5177(83)90045-6.

[14] Hyejin Yoon and Edward J. Malecki. "Cartoon Planet: Worlds of Production and Global Production Networks in the Animation Industry". In: *OUP Academic* (2009). URL: https://academic.oup.com/icc/article/19/1/239/666922#11645160.