**University of Oklahoma**

Logo, company name

Description automatically generated

*Motto: To “sea” the world through the lens of data*

Team Sea World

Andrew Clements, Kolten Kemper, Lorenzo Rocchi, and Vishal Reddy

Business Data Analysis – MIT 5682

Professor Swetha Siripurapu

May 4th, 2023

# **Executive Summary**

SeaWorld Orlando is facing impending financial troubles, likely from improper revenue management. There was a previous lack of importance assigned to additional revenue streams in the form of meal passes, parking passes, and gift shop sales. The decreased attendance in the winter months is addressable since the park’s location is unaffected by cold winter temperatures, presenting an opportunity for revenue growth.

The wide yet unfocused net cast by SeaWorld’s marketing strategy is also of concern. From contribution analysis, we see that most states have a uniform distribution of ticket sales, with the exception of neighboring states. Targeting such a mass market has led to an underperformance of sales. Instead, SeaWorld should focus on existing regional states with higher sales, as well as states in the vicinity of the regional states. that incur fewer travel expenses and are more likely to attend the park.

Finally, the high correlation between youth ticket revenue and additional revenue streams suggests more bundle deals with additional parking passes, meal passes, and gift shop merchandise should be explored. Likewise, a Pareto analysis shows that approximately 60.8% of customers contribute 80% of the park’s revenue. Therefore, a more concentrated strategy could tap into a more profitable segment.

From this, this report details further analysis, including key metric analysis, comparison analysis, trend analysis, ranking analysis, variance analysis, contribution analysis, frequency analysis, correlation analysis, and Pareto analysis. Total cost of this comprehensive analysis was $540 and 1,295 minutes.

Contents

[Executive Summary 2](#_Toc1263679557)

[Get to Know the Team: Team SeaWorld 4](#_Toc2089831333)

[All About Data 5](#_Toc1360449143)

[Explain what you are trying to solve 5](#_Toc207404237)

[What are the different types of Data 5](#_Toc35810420)

[ETL 6](#_Toc668889136)

[What is ETL? Why is it necessary? 6](#_Toc1861275929)

[Why is the Process of Connecting And Analyzing Data 6](#_Toc294371006)

[Why Analyze Data 6](#_Toc763506061)

[Big Data 7](#_Toc1124386646)

[How many ways can we Get Data into Excel 8](#_Toc1026510317)

[What are the Rules for Preparing Data 8](#_Toc909945988)

[What data/usecase did you choose for your Analysis 8](#_Toc2144502907)

[How did you prepare your data for Analysis 8](#_Toc1815557918)

[Cleaning Data 8](#_Toc837569884)

[Key Metric Analysis 9](#_Toc1310684572)

[Add Key Metric Analysis that you did for your project 9](#_Toc2047964267)

[Pivot Tables 11](#_Toc2054753272)

[Comparison Analysis and Importance 12](#_Toc1026963986)

[Comparison Analysis with Example 12](#_Toc1426479200)

[Trend Analysis 14](#_Toc195924308)

[Implement Trend Analysis with example 14](#_Toc40360443)

[Ranking Analysis 17](#_Toc1926540772)

[Interactive Dashboards 19](#_Toc471516903)

[Variance Analysis 20](#_Toc1542795155)

[Contribution Analysis 21](#_Toc1734007652)

[Frequency Analysis 22](#_Toc529272186)

[Correlation Analysis 24](#_Toc1121791212)

[Pareto Analysis 25](#_Toc532121987)

[Challenges Faced/Addressed During Implementation 29](#_Toc1589249461)

[Strengths and Weaknesses Encountered During Implementation 29](#_Toc535718178)

[What We Learned Throughout This Process 29](#_Toc2140660218)

[Appendix 31](#_Toc1040026080)

[Project Management 31](#_Toc343888099)

# **Get to Know the Team: Team SeaWorld**

|  |  |
| --- | --- |
| A person in a suit smiling  Description automatically generated with medium confidence | A person smiling for the camera  Description automatically generated with medium confidence |
| **Kolten Kemper** | **Andrew Clements** |
| MBA/MS MIT Dual-Degree | MBA/MS MIT Dual-Degree |
| 1st Year | Professional |
| IT Advisory and Actuary | Purchasing and Supply Chain |
| Undergrad in Finance and Accounting from OU. | Undergrad in Geology from Washington & Lee. Masters in Geophysics from Georgia. |

|  |  |
| --- | --- |
| A person smiling for the camera  Description automatically generated with medium confidence | A person wearing glasses and a suit  Description automatically generated with medium confidence |
| **Vishal Reddy Musku** | **Lorenzo Rocchi** |
| MS MIT | MBA/MS MIT Dual-Degree |
| 1st Year | 2nd Year |
| Accenture Technologies | Gnowbe Inc. and Auerbach Grayson & Co. |
| Bachelor of Technology in Computer Science from JB Institute of Engineering & Technology | Bachelor of Business Administration in Marketing and International Business with a concentration in Japanese from OU. |

# **All About Data**

A datum is a singular piece of information. The plural, data, is a collection of pieces of information. Data is useful for creating greater and measurable understanding of a scenario, as well as creating the ability to make informed and meaningful determinations, conclusions, and decisions. From this, data is an essential tool in the world of business. However, it must be treated with intentionality and purpose, as it is susceptible to manipulation.

## **Explain what you are trying to solve**

We seek to find answers to various business questions involving different metrics in order to improve the profitability and operations at SeaWorld’s Orlando park, as well as discover opportunities that may have previously lacked optimization. With the recent COVID-19 pandemic negatively affecting public activities, as well as overall declining popularity in the SeaWorld theme parks, SeaWorld is plagued by a problem of stagnant, suboptimal revenue. Thus, through various types of business data analysis, including key metric analysis, comparison analysis, trend analysis, ranking analysis, variance analysis, contribution analysis, frequency analysis, correlation analysis, and pareto analysis, we will try to solve the underlying questions posed as to how SeaWorld can increase its Orlando park’s revenue.

## **What are the different types of Data**

Data largely falls into two categories: qualitative and quantitative. These can be further decomposed into various subcategories. Qualitative data can be divided into nominal data like names and ordinal data, with relative rankings or positions. By contrast, quantitative data breaks down into discrete data which cannot be divided, such as a count of inventory items in an ERP, or continuous data that divisible such as lengths of time to complete a product. Finally, continuous quantitative data can be grouped into interval data, such as temperature, and ratio data where there is a true and meaningful zero such as weight.

Additionally, there are status and event forms of data. Status forms of data describe the current state of something. For example, it could be the current inventory of an item in a department store or warehouse. On the contrary, event data describes an instance in time that had a sudden occurrence. For example, a transaction id, or a purchase date would be an event for data, as it describes the point in time of a particular event and will never change. Furthermore, there is transient and periodic data. Transient data are common in operational systems and are subject to change and not preserved. Period data is dissimilar, as it typically resides in a data warehouse and is never altered.

Finally, there is unstructured and structured data. Structured data id clearly defined, such as a number or word, also with a definite format. Unstructured data does not follow a specific format. These can include images, pictures, social media posts, and more.

# **ETL**

ETL stands for Extract, Transform and Load. An ETL tool extracts the data from different RDBMS source systems, transforms the data like applying calculations, concatenate, and more. and then loads the data to different systems to perform analysis.

Some of the assumptions of an ETL include extraction and automation. First, data is extracted from one or more source systems. The extracted data may require some form of transformation or cleansing to ensure its quality, accuracy, and consistency. Second, ETL tools are used to automate the extraction, transformation, and loading of data into a target system. In the case of SeaWorld, we are assuming it has been extracted from multiple sources and concatenated properly, possibly by automation.

## **What is ETL? Why is it necessary?**

ETL stands for extract, transform, load. It is the process by which data is taken from multiple sources and consolidated into a single repository. It’s necessary because data needs to be prepared, combined, and loaded to be analyzed and modeled to determine insights. Being in a centralized location will increase the quality, consistency, and speed of the business operations. It is important that we use one in the case for SeaWorld since it is a large firm with many different operations. For example, we need to connect the ticket booth data with the online ticket sales data, which will make it easier to analyze ticket sales.

## **Why is the Process of Connecting And Analyzing Data**

Connecting data allows you to analyze data in multiple positions that continuously updates from other positions in a spreadsheet. The process consists of analyzing, transforming, loading, and analyzing data. First, extracting the data simply involves getting files from the place needed. Here, you will likely contact someone working in that business segment, or someone from IT, Second, when you have the data, you will transform it into the desired format. Third, loading the data into the desired place will occur to have it be centralized. Finally, analysis will be performed, in which various analyses such as key metric and contribution will occur, and insights will be uncovered.

## **Why Analyze Data**

Analyzing data allows you to move from speculation and assumptions about business questions to improve operations, financial performance, and decision-making; such questions often involve significant amounts of money with potentially large impacts on customer experiences and park profitability, more soundly with evidence. Therefore, it is essential to use analyses to make evidence-based decisions that will remove as much behavioral bias as possible. This aids in our goal of making sound, data-driven results to improve SeaWorld Orlando’s operations.

# **Big Data**

Big Data is data of considerable size that requires a storage system, such as Hadoop. Big data is usually characterized by those who have to innovate it to analyze it. This phenomenon is changing the world in the way we conduct business in that companies have enormous amount of data they must protect. These include significant costs of maintaining such data through cybersecurity. Additionally, foreign governments, such as the United Kingdon, have begun to support projects in big data, investing significant capital.

Businesses are using transaction data for recording sales, web and server logs for webpage traffic/search info, machines/sensor data, email and documents, social media to stay connected to the customer base, geospatial for satellite images, digital images, video for security, and audio. This has greatly benefitted businesses in the form of increasing reporting accuracy, optimization of operations, faster analysis, improved management, compliances, product innovation, and improved customer marketing.

Big data is primarily differentiated from regular data by the “three Vs” of big data. The first V pertains to volume, or size of the data. There is not an explicit number of bytes in which data becomes “big,” though it can largely be perceived as sufficiently large to where the data cannot be analyzed, collected, and stored by traditional means without the use of a data warehouse or data lake. The second V, velocity, refers specifically to the speed to collect, process, and analyze the data. Big data is generally characterized by a large, constant flow of data such as photos uploaded to Facebook or transactions at point-of-sale terminals. Finally, big data is characterized by its variety. Unlike traditional sources of data, big data is a collection of data across many different data types often unrelated to one another. Additionally, the data varies in type between structured data, such as transaction data versus unstructured data such as photo and video.

## **How many ways can we Get Data into Excel**

We can import data from file, database, Azure, Power Platform, online services, and other sources. On top of this, there are many subsets of ways you can get data into Excel. For example, getting data from a file could be a plethora of file types such as a csv, txt, json, and more.

## **What are the Rules for Preparing Data**

Some of the rules for preparing data are as follows. First, when defining column names, ensure they accurately and meaningfully represent the data they contain. Second, each row in the dataset represents a single record of data, containing all the relevant information for that particular entry. Third, each row in the dataset represents a single record of data, containing all the relevant information for that particular entry. Fourth, the dataset should only contain raw data and should not include any subtotals or grand totals. Fifth, ensure that the data types of each column are accurate and correctly reflect the nature of the data, whether it is text, date, numeric, currency, or other data types. Sixth, verify that text fields are left-aligned within their cells, while numeric data is right-aligned for consistent and easy-to-read formatting.

# **What data/usecase did you choose for your Analysis**

For the Director of Operations of the SeaWorld Park in Orlando, Florida. We are analyzing the level of ticket sales, ticket bundle performances, performance of meal passes, parking passes, and ticket prices in relation to the time of year/day of the week from a management perspective. We can analyze the data to compare different ticketing bundles and revenues based on timing factors, as well as perform a trend analysis to project future performance. From this, this analysis will aid in your park’s overall efficiency in forecasting demand and resource allocation, as well as increasing revenue.

# **How did you prepare your data for Analysis**

We prepared our data for analysis by using generatedata.com in which we selected salient factors that would be pertinent for a revenue analysis of SeaWorld. We included a focus on revenue, tickets, and passes that would be used to analyze the inadequacies of revenue generation. We prepared total revenue by manually calculating it by randomly generating a random amount of tickets purchased, along with multiplying that number with the posted SeaWorld ticket prices. Then, we imported the CSV file into Excel for analysis.

## **Cleaning Data**

Cleaning is a process critical to working with data. It serves as the foundation for rendering data uniform and interpretable by the tools that will be used. For instance, data that has not been cleaned will be unreadable to excel and not be able to be made effective use of in excel. Because we had strict control of the data generation output, minimal cleaning was required. The main process of cleaning largely revolved around adding calculated columns to better summarize data such as the total number of tickets in a transaction and total revenue in a transaction or reflecting a reduction in ticket price for online purchases.

# **Key Metric Analysis**

Key metric analysis is a way of looking at important numbers that show how well a business or organization is doing. These metrics or indicators might include things like sales and other variables in the form of sum totals, counts, maximums, minimums, or averages. By analyzing these metrics, businesses can figure out what's working well and what needs to be improved. They are used to improving the performance of businesses by taking better decisions.

## **Add Key Metric Analysis that you did for your project**

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the total quantity (count) of ticket sales? |  |
| 2 | What is the max quantity of ticket sales in a transaction? |  |
| 3 | What is the min quantity of ticket sales in a transaction? |  |
| 4 | What is the average quantity of ticket sales per transaction? |  |
| 5 | What is the total ticket revenue? |  |
| 6 | What is the max total ticket revenue in a transaction? |  |
| 7 | What is the min total ticket revenue in a transaction? |  |
| 8 | What is the average ticket revenue per transaction? |  |
| 9 | What is the total quantity (count) of parking passes sold? |  |
| 10 | What is the total revenue of the parking passes? |  |
| 11 | What is the total quantity (count) of meal passes sold? |  |
| 12 | What is the max total quantity of meal passes sold? |  |
| 13 | What is the min total quantity of meal passes sold? |  |
| 14 | What is the average total quantity of meal passes sold? |  |
| 15 | What is the total revenue of the meal passes? |  |
| 16 | What is the max total revenue of the meal passes? |  |
| 17 | What is the min total revenue of the meal passes? |  |
| 18 | What is the average total revenue of the meal passes? |  |
| 19 | What is the total revenue of gift shop transactions? |  |
| 20 | What is the max revenue of gift shop transactions? |  |
| 21 | What is the min revenue of gift shop transactions? |  |
| 22 | What is the average revenue of gift shop transactions? |  |
| 23 | What is the average total revenue of concessions transactions? |  |
| 24 | What is the max revenue of concessions transactions? |  |
| 25 | What is the min revenue of concessions transactions? |  |
| 26 | What is the average revenue of concessions transactions? |  |

In analyzing these ten questions, we obtained summary statistics that defined SeaWorld’s key performance indicators. In this instance, we wanted to provide an overview of overarching statistics that describe the ticket revenue segment, parking pass segment, meal pass segment, concessions segment, and gift shop segment.

## **Pivot Tables**

A pivot table is a data summarization tool used to analyze, summarize, and transform large amounts of data. Pivot tables are especially useful with large amounts of data that would be time-consuming to calculate by hand. The main advantage is their dynamic nature that allows them to quickly be manipulated. This can be seen in their ability to read data sets quickly and extract header column names that the data analyst can then drag in drop into various row and header configurations, and also change the number format the data is being presented in.

A picture containing text, screenshot, font, number

Description automatically generated

# **Comparison Analysis and Importance**

Comparison analysis is a process of comparing two or more things to identify similarities and differences between them. It helps in identifying trends and patterns, evaluating performance, and making informed decisions. It is useful to compare the performance of similar components. Our data is mainly focused on revenue analysis, so it will analyze various sources of revenue and how they compare. Since SeaWorld is a destination service, time of the year is important, making comparison of time with revenue focal point.

## **Comparison Analysis with Example**

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| **1** | How does the most common order size compare with each other? |  |
| **2** | How does revenue compare between various sources (i.e., concessions, merch, tickets, etc.)? |  |
| **3** | How does total revenue compare by month? |  |
| **4** | How does the amount of standard day tickets compare to the amount of season passes sold? |  |
| **5** | How does revenue at the concession stands compare in terms of meal passes and spot purchases? |  |
| **6** | How do the bundling options compare for each transaction? |  |
| **7** | How do the ticket sales compare by day of the week? |  |
| **8** | How does the purchase medium compare for total transactions and total number of tickets sold (online vs. in-person)? |  |
| **9** | What is the comparison of total ticket revenue based on ticket types (adult, child, senior, etc.)? |  |
| **10** | What is the comparison of number of tickets sold based on marketing types? |  |

## **Trend Analysis**

Trend analysis is the examination of data over time to identify patterns or trends. It involves analyzing historical data to uncover information to predict the future. We used trend analysis to determine during what time of the year the ticket sales increase based on the historical sales information. SeaWorld can then use this to assess when to offer promotions in order to take advantage of such trends, or smooth revenue by focusing on down periods.

## **Implement Trend Analysis with example**

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the trend of ticket sales by year and month by ticket type? |  |
| 2 | What is the seasonality of ticket sales by month? |  |
| 3 | What is the trend of parking pass sales by year and month? |  |
| 4 | What is the trend of meals pass sales by year and month? |  |
| 5 | What is the trend of gift shop sales by year and month? |  |
| 6 | What is the trend of total ticket quantity per transaction by year and month? |  |
| 7 | What is the seasonality of ticket quantity per transaction? |  |
| 8 | What is the seasonality of online vs in-person ticket purchases by year and month? |  |

From the trend analysis, we see the overarching trends stay fairly stable for the majority of the business questions posed. However, the seasonality of ticket sales by month displays the presence of increased activity in the summer months. This could lead to management focusing on resource allocation for the summer months to capitalize on this increased presence. Additionally, in question 7, there seems to be a spike in the summer months of the average party size, indicating families usually attend the park in the summer. This can also lead SeaWorld’s management to push more family promotions in the summer months.

# **Ranking Analysis**

Ranking analysis is a technique used to compare and evaluate the relative performance of different entities based on their rankings. We used the ranking analysis to find out the states with the highest customers, and used to find out the months with highest gift shop revenues, etc.

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the ranking of ticket sales by month? |  |
| 2 | What is the ranking of total revenue by month? |  |
| 3 | What is the ranking of parking passes sold by month? |  |
| 4 | What is the ranking of meal passes sold by month? |  |
| 5 | What is the ranking of most popular ticket types by month (child adult)? |  |
| 6 | What is the ranking of online ticket transactions by month? |  |
| 7 | What is the ranking of park attendee’s origin states? |  |
| 8 | What is the ranking of marketing types among customers for total revenue? |  |
| 9 | What is the ranking of park attendance by year? |  |
| 10 | What is the ranking of gift shop revenue by month? |  |
| 11 | What are the top 12 states by park attendee? |  |

The ranking analysis was focused on monthly rankings for total revenues, as well as the ancillary arms that make up additional revenue sources. The ranking analysis shows summer months of May through July are consistently lucrative on numerous fronts. Surprisingly, the ranking of park attendance by orgin state is more uniform than expected. To this end, management should consider increasing their current footprint in the surrounding Florida area.

## **Interactive Dashboards**

Interactive dashboards are interfaces that allow users to visualize and interact with data in real-time easily and more effectively. They allow customizable views and collection of multiple data visuals for better consumption of data by the users. For this, we compile all metrics and analysis we deemed important to monitor to have it in a central location that allows a non-expert to better monitor SeaWorld’s operations.



## **Variance Analysis**

Variance analysis is a technique used to identify the difference of an entity/multiple entities with respect to another entity. It assists largely with comparative measures and deducing changes or differences. This can be used to detect unexpected changes, or for budget targets that indicate how well a firm adhers to a goal.

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the variance in park attendance from previous months? |  |
| 2 | What is the variance in total revenue from previous months? |  |
| 3 | What is the variance in gift shop revenue from the previous month? |  |
| 4 | What is the variance in parking pass sales from previous months? |  |
| 5 | What is the variance in meal pass sales from last month? |  |

## **Contribution Analysis**

Contribution analysis is a technique used in data analysis to identify the relative contributions of different factors to a particular outcome or result. This is mainly how much a particular factor affects another total factor. For example, how much of total revenue is contributed by a gift shop revenue.

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the contribution percentage of revenue from each revenue source? |  |
| 2 | What is the contribution of revenue from each ticket type? |  |
| 3 | What is the contribution of total ticket sales by attendee state of origin? |  |
| 4 | What is the contribution of revenue from each month? |  |
| 5 | What is the contribution of each marketing type to total revenue per year? |  |

## **Frequency Analysis**

Frequency analysis is a statistical technique used to study the frequency of occurrence of certain events or values in a dataset. From this, we can see how likely something is to happen and assess whether or not there is statistical significance for the finding. From this, we can assess whether to make a certain business decision/recommendation.

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the frequency of customers who spend between 0 and 400 dollars on meal passes, in intervals of 50? |  |
| 2 | What is the frequency of customers who spent between 0 and 3500 dollars in total, in intervals of 500? |  |
| 3 | What is the frequency of customers who spent between 0 and 2500+ dollars in tickets, in intervals of 250? |  |
| 4 | What is the frequency of customers who spent between 0 and 200 dollars in concessions, in intervals of 25? Also, by whether or not they bought a meal pass. |  |
| 5 | What is the frequency of customers who spent between 0 and 200 dollars in the gift shop, in intervals of 25? |  |

## **Correlation Analysis**

Correlation analysis is a technique used to measure the strength and direction of the relationship between two or more variables. If two variables are positively correlated, if one moves in one direction, the other variable will move in the same direction. If two variables are negatively correlated, the variables will move in the opposite direction.

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What is the correlation between season pass sales and ticket revenue by month? |  |
| 2 | What is the correlation between youth ticket revenue and adult ticket revenue by month? |  |
| 3 | What is the correlation between meal pass revenue and concessions revenue by months? |  |
| 4 | What is the correlation between the number of Parking passes sold and number of season passes sold by month? |  |
| 5 | What is the correlation between total concessions revenue and total youth ticket revenue? |  |

## **Pareto Analysis**

Pareto analysis (80/20 rule), is a technique used to identify the most significant factors or problems that contribute to a particular outcome or result. The pareto principle posits that 80% of outcomes are put forth by 20% of the inputs. Mainly, we found most of the time, 55% - 60% of the customers contribute 80% of the revenue.

|  |  |  |
| --- | --- | --- |
| **Query #** | **Question** | **Output** |
| 1 | What percentage of Marketing types contribute to 80% of the total ticket revenue? |  |
| 2 | What percentage of States contribute to 80% of the total tickets sold? |  |
| 3 | What percentage of customers contribute to 80% of the total revenue? |  |
| 4 | What percentage of customers contribute 80% of the concession revenue? |  |
| 5 | What percentage of customers contribute 80% of the Gift shop revenue? |  |

## **Challenges Faced/Addressed During Implementation**

1. Creating a representative dataset with workable data. We randomly generated data and ensured it was properly formatted so we could do various analyses.
2. Choosing the most important/right questions from the SeaWorld Park’s management perspective. We, as a team, discussed all the different points and aligned our thoughts.
3. Creating the calculations during the data cleaning process to ensure that we can accurately tabulate measurements such as revenue for each source and ensuring that Boolean logic was implemented correctly.

## **Strengths and Weaknesses Encountered During Implementation**

The diverse backgrounds of all the team members were a huge advantage while working on the project as there were multiple views that helped others to look at things from a fresh perspective. Additionally, another strength was that each member was strongly individually and intrinsically motivated meaning that there was little need to coax members into performing the work and putting effort toward the project. Finally, each team member was willing to explain the rationale and process behind their suggestions thus minimizing confusion and assisting to ensure that each member understood the process.

The main weakness was largely a result of inexperience. Nobody on the team came into the project necessarily as an experienced expert in using pivot tables, but this was remedied through the length of the course. Another weakness was that communication was a bit difficult due to differing time schedules, but this was overcome with the strong motivation to take ownership of the work by each member as well as being well suited to autonomous work.

# **What We Learned Throughout This Process**

In this process, we learned many useful tools about the individual processes associated with business data analysis. While the class was beneficial, we all agree that we learned the most when we were creating ideas together in terms of implementing the analysis with the data we created. We learned that working on a project from start to finish is the best way to learn the skills we each took away from this experience.

|  |  |
| --- | --- |
| **Member Name:** | **What you learned:** |
| Kolten Kemper | I learned how to apply Pareto analysis, which is a widely effective tool in the growing world of data now. Knowing how to do a sought, complex task is widely useful. I also learned the importance of using slicers, which make it fast and efficient to display different scenarios for the data. This would have been very beneficial for me at my previous job. |
| Lorenzo Rocchi | I learned how to operate a great variety of pivot table functionalities as well as how to use previously known techniques to gain new insights, such as the Pareto analysis. The true power of pivot tables made me realize that data analysis is dynamic, more so than I originally thought. I plan to use this knowledge to work faster in Excel and use pivot tables to analyze problems quickly. For example, I will do this instead of the typical “count” function I used previously to count iterations of various functions. This will make my life much easier. |
| Vishal Reddy Musku | I learned about different types of analyses, what analysis should be used when and how to use the right data visuals to show the results. I learned that column charts are best for showing differences in quantities, pie charts are best for showing contribution, and line charts are best for showing trends. This will be very useful in my career for making effective data visuals. I also learned how quickly they can be made. Even if I do not need a visual for a final report, I can quickly make one to analyze relevant information at a quick glance. |
| Andrew Clements | I learned how to clean data and how to quickly analyze it in different meaningful ways, especially with visuals. This taught me the importance of making sure raw data remains clean. I originally thought it was important to leave data untouched, but I now realize there is a difference between manipulation and “cleaning.” I will use this to ensure all future dealings with data do not have errors, and my spreadsheets remain clean. |

The project management tool proved to be very useful. It kept us all accountable and ready to work together because we all wanted to be viably always helping others. We all enjoyed it and would likely use it again.

# **Appendix**

## **Project Management**

Team Name: Team SeaWorld

Logo:   
Logo, company name

Description automatically generated

Team Motto: To “Sea” the World through the lens of data

**Team Members**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Email | Phone | Strengths | Availability to Meet |
| Andrew Clements | Andrewgclements@ou.edu | (405)343-3625 | Academic writing, attention to details | Most days, but live in OKC |
| Kolten Kemper | koltenkemper@ou.edu | (918) 629-2618 | Attention to detail, quality control, math, diligence, communication | Every day except Wednesday |
| Lorenzo Rocchi | Lorenzo.b.rocchi-1@ou.edu | (214) 592-2134 | Long-form academic writing | Every day except Tuesday from 1pm to 5pm, can adjust with advance notice |
| Vishal Reddy Musku | Vishal.Reddy.Musku-1@ou.edu | (980)318-1432 | Attention to detail, academic writing, math | Tuesday, Friday, Saturday, and Sunday |

**Team Expectations for the confidential peer evaluation:**

Punctuality, good communication, good effort, good teamwork.

**The behavior for which points will be deducted on the confidential peer evaluation:**

Not doing one’s fair share of work, consistently missing meetings without notice, or lack of cooperation

