Data Analytics and Organisational Decision Making: Group Coursework

MANG 6526

Student ID#: 31302815, 33368961, 33083143

Word Count: 6039

# Part A

## 1. Identify and discuss critical problems related to making effective organisational decision

In case study about *The Cage* by *Perdana et al (2020)*, there are a variety of problems related to effective organisational decision making. The following section will use established using decision making theories and research to identify and explain decision making problems in *The Cage*.

### 1.1 Rational Decision-Making Model

*The Cage* case study displays decision making problems relative to the "Rational Decision-Making Model" as stated in an article by *Mustafa & Kingston (2014)*. This theory suggests that the individual making decisions understands the problem, creates and reviews all possible solutions and alternatives, and identifies goals from which objectives can be made (Mustafa & Kingston., 2014). When comparing the Rational Decision-Making Model to *The Cage*, their decision making process is unparallel. The owner Sanjay understands *The Cage*'s problem in that their current use technology is not working, competition in their market is picking up, and they are not making use of all their data (Perdana et al., 2020). However, when identifying solutions, Sanjay only proposes advanced Data Analytics (DA) because he is convinced that it will benefit the company, he does not consider other options nor outlines how exactly he will use advanced DA to benefit the company (Perdana et al., 2020).

### 1.2 Prospect Theory

Decision making problems are also apparent in *The Cage* in comparison to "Prospect Theory" as stated in *Mustafa & Kingston (2014)*. In prospect theory, decision makers take risks in decisions when they are facing an undesirable outcome and are less risky in decisions when they are considering gaining an advantage (Mustafa & Kingston., 2014). *The Cage* is doing the complete opposite of Prospect Theory as they are putting more risk into a decision— the implementation of advanced DA—for the believed benefit of the company (Perdana et al., 2020). *The Cage* is not facing an undesirable outcome, so in regard to prospect theory, should not be so risky when making decisions. They do behave not risk aversive in their approach to deciding on advanced DA, which they appear to be fully invested towards (Perdana et al., 2020).

### 1.3 Group Think

The problematic decision making in *The Cage* can also be attributed to "Group Think" as stated in *Buchanan & O'Connell (2006)*. "Group Think" is a poor quality decision making concept referring to when people within a group are so focused on the cohesiveness of the group, that it leads to unanimity of group decisions without consideration for other possible options (Buchanan & O'Connell., 2006). "Group Think" is reflected in *The Cage*'s decision making process regarding DA (Perdana et al., 2020). The "Group" making the decision in *The Cage* are Sanjay (the owner) and Anthony (IT Advisor). Sanjay is the one who initiates talk about implementing advanced DA, who's excitement about it also excites Anthony (Perdana et al., 2020). During a discussion, Sanjay questions whether they can make use of all company data—referring to whether advanced DA is capable of this (Perdana et al., 2020). Anthony's immediate response is filled with excitement as he agrees that they can do it, but also acknowledges that it will be difficult (Perdana et al., 2020). This interaction demonstrates "Group Think" because Anthony agrees with Sanjay in a way that he wants to maintain the consensus that advanced DA is the right decision, but he also acknowledges the associated difficulties (Perdana et al., 2020). The fact that Anthony, being the IT advisor, acknowledges the difficulties and does not present any alternatives, but agrees with Sanjay, demonstrates "Group Think" (Perdana et al., 2020).

### 1.4 Political View

Similar to "Group Think", The "political view" theory as stated in *Turpin & Marais (2004)* highlights similar decision making problems in *The Cage*. The "political view" is a theory where decisions are made based on personal agendas (Turpin & Marais., 2004). In the political view decision model, parties involved have competing views and goals in an ongoing decision making process (Turpin & Marais., 2004). In comparison to *The Cage*, the decision makers do not follow a political view model as there are only two individuals involved in the decision making, Sanjay and Anthony, and they both hold the same, non-competing views in that they both believe advanced DA is what the company needs (Perdana et al., 2020). Anthony is the other party in the decision making process in *The Cage*, and as the main IT Advisor, not even he suggests any alternatives to DA (Perdana et al., 2020).

### 1.5 Gut decision making

Gut decision making is another low-quality decision making technique addressed in *Buchanan & O'Connell (2006)* that is reflected in *The Cage.* Gut decision making is based on the confidence of the one making decisions (Buchanan & O'Connell., 2006). They are not heavily based on evidence and are made in situations of crisis where there is limited time to consider outcomes (Buchanan & O'Connell., 2006). Gut decision making is reflected when Sanjay from *The Cage* is deciding whether to implement advanced DA (Perdana et al., 2020). *The Cage* is in a moment of crisis as their current use of technology is not working and competition is rising in their market (Perdana et al., 2020). This moment of crisis provokes Sanjay to make the decision to pursue advanced DA (Perdana et al., 2020). Sanjay claims that advanced DA will improve decision making, strengthen government ties, and reveal hidden market segments, but has no evidence to support his claims, he is only convinced that it will accomplish these tasks (Perdana et al., 2020).

### 1.6 Political-behavioural decision approach

Finally, problems in *The Cage*'s decision making process can be attributed to the "political-behavioural decision approach" described in *Ilori & Irefin (1997)*. This theory describes when decision makers of an organisation face pressure from external stakeholders who can affect purpose of an organisation and cause the organisation to make decisions to satisfy these stakeholders (Ilori & Irefin., 1997). In relation to *The Cage*, one of the main reasons as to why Sanjay wants to use advanced DA, is so *The Cage* can provide the authorities with factual data so that they have a reason to keep supporting the business (Perdana et al., 2020). The government of Singapore is a stakeholder in *The Cage*, as their policy of encouraging citizens to partake in exercise and sport, drives business to *The Cage* (Perdana et al., 2020). Hence, *The Cage* is pressured to implement advanced DA to give factual evidence to the government, so they continue to provide support. This approach limits *The Cage's* ability to consider other options.

## 2. Critically discuss how should The Cage effectively use social media as a strategy in data analytics.

Social media has provided rich data sources to study people's behaviour in nature and has become one of the critical factors for business excellence since Web 2.0 (Di Minin et al., 2015). Based on a report by Bain & Company, companies that adopted big DA were five times faster in decision making than their competitors and twice as likely to be in the top quartile of financial performance within the industry (Pearson & Wegener., 2013). Social media is not only a platform that offers businesses a place to launch their products or services, but it acts as a dataset that helps them to analyse and incorporate customer insights into their decision-making processes. Moreover, the emergence of social media sites has made customer power even stronger (Labrecque et al., 2013). Companies can better understand their target audience and develop strategies analysing end customer comments from social media. As Sanjay mentioned, they would like to use data to influence authorities (Perdana et al., 2020). Knowing the relationship between social media and the data on hand could predict future sales and peaks

It is important to note that retrieving usable data from social media is not an easy task due to its large volume and its variety of contexts (i.e. different communication patterns) (Schreck and Keim., 2013).

### 2.1 Social Media Challenges for The Cage

In the present, The Cage only possesses corporate website and Facebook as their social media tools. However, it is necessary to expand social media variety since social media communications impact brand visibility as an electronic word-of mouth(Casaló et al., 2010). Moreover, customers' online experience which involves richness of content, service, and product, and customized interaction is believed to influence business promotion (Goswami et al., 2013). Some specific ways have also been suggested in the early research by *Drèze and Zufryden (2004)* to enhance the brand visibility through key words and link positions. Thus, to improve brand visibility as Sanjay requested, handling various social media accounts is the most critical issue regarding social media marketing.

In terms of DA, the first challenge is related to multimedia sharing on social media platforms as it increases different forms of information, leading to complexity of data processing (Moise et al., 2013). For example, scientists can differentiate a post's intention through textual posts using text-mining techniques (Adamopoulos et al., 2018). Other platforms using different media, such as YouTube, data scientists should adopt other methods to transform media (i.e. videos) to aggregate all the information from other platforms.

Secondly, the limitation of access to user data and change in user behaviour leads to data variance (August et al., 2020). The digital divide excludes some people from data collection, causing organisations to neglect these potential customers (McKenna et al., 2017). In this case, social media cannot be a resource if The Cage wants to expand or enhance their services to specific customers. Apart from the technical constraints to accessing data, the scientists using social media data also need a certain level of access on the platform (McKenna et al., 2017). For example, on platforms such as Facebook, some individual(s) need permission to view specific groups and their information. Scientists need to be able to navigate the platform and access specific areas. Social media is prone to anonymous and fake users that can make it difficult to distinguish between valid or false information (McKenna et al., 2017). Since everyone has the right to express themselves on social media, there is less control of information, and data is at risk of being unauthentic (McKenna et al., 2017). As a result, data crawled from social media might not prove useful to scientists working with it.

## 2.2 Recommendation for The Cage

In order to mitigate the challenges associated with social media mentioned earlier, recommendations can be made regarding Data mining and enhancement of practice.

### 2.2.1 Employ Data Mining for Decision Making

To resolve the issue of data integrity, The Cage should adopt unsupervised data-learning methods to organize documents via clustering and filtering information according to user preferences (Banerjee & Basu, 2007). In particular, sentiment analysis is a method that can extract opinion-related views and determine people's attitudes either negatively or positively towards a specific topic from unstructured text to further assist in business choices (Jindal and Aaron, 2021).

In addition, the trends observed from the data can assist The Cage in maintaining its supply-demand cycle in the competitive markets (Patil et al., 2017). The Cage can categorize customer requirements and customize its services according to different segments. It is reported that 80% of consumers are likely to do business with companies that provide personalised services (Epsilon, 2018). Therefore, The Cage can strategize its business practices by fulfilling customer acquisitions, obtaining higher retention, and reducing costs for customer care (Mäntylä, 2017).

Lastly, the data-driven results can be applied to the decision making process in The Cage. *Chan et al. (2016)* exploited social media data in New Product Development (NPD) and Multi-Criteria Decision-Analysis (MCDA) to leverage operation management decision-making based on customers' voices. Understanding customers' requirements and preferences are key factors of successful NPD (Katila & Ahuja., 2002, Piller and Walcher., 2006, and Von Hippel., 1986).

Also, the social media expenditures can be justified by NPD, which helps companies stand out from a competitive standpoint and become financially sustainable (Henard & Szymanski., 2001, Krish-nan & Ulrich., 2001).

#### 2.2.2 Enhance Practice

In recent years, data gathering systems such as Supermetrics.com have assisted marketers in monitoring and aggregating the advertising effectiveness from different social media platforms (Supermetrics., 2022). The Cage should choose a system—as aforementioned—that allows real-time access to combine data from different platforms and examine the marketing effectiveness. As mentioned previously, scientists need to investigate the website to understand the data structure better, so we suggest that The Cage obtain permission from platform gatekeepers or socialize among users within platforms.

Social media promotion plays a diverse role in terms of DA. Using results from DA, The Cage can utilise them to enhance their hosting of events and differentiate their customers on various platforms. For instance, on Instagram, The Cage can hold interactive events such as quizzes or hashtags to attract younger audiences and establish a ludic image. Customers are encouraged to consistently engage with The Cage and develop a track record of being there with fresh innovation when new needs arise. Simultaneously, The Cage can obtain brand visibility through interaction between social media and followers.

In some way, social media and DA benefit one another. Analysing data on social media allows for companies, including The Cage, to identify their target customers (Moe & Schweidel., 2017). The more activity on social media, the more data generated. In the case of The Cage, they can conduct quick surveys on different social media platforms to gain knowledge on people's exercising frequency.

## 3. What ethical concerns do you think The Cage would face in terms of using data analytics? What recommendations would you give the company to address these concerns?

### 3.1

In the case study, The Cage relied on a book-and-pay system for data collection. After obtaining customer data by telephone call booking, online booking, and on-site booking, The Cage would use the data to open hidden market segments and maintain their competitive advantage. Although The Cage tried to leverage data for analysis, they acknowledge that they have a lack of expertise, which puts focus on ethical concerns regarding DA (Perdana et al., 2020). Ethical concerns can be categorized into several aspects.

#### 3.1.1 Privacy Concern

The privacy concern present in The Cage is that they collect customer data without customers completing the Personal Data Protection Act (PDPA) consent form—a mandatory requirement for all businesses in Singapore (Perdana et al., 2020 & Personal Data Protection Commission., 2022). This is a concern because they do this even after promising that customer data is secure (Perdana et al., 2020). For example, when customers book online, they non-consensually collect the customer's name, address, and phone number (Perdana et al., 2020). *White & Ariyachandra (2016)* suggest that when personal information is shared with others, especially without their agreement, many people are threatened with losing privacy, which eventually leads to customer resistance. The Cage violated the PDPA, meaning they cannot guarantee that customer rights are fully protected.

#### 3.1.2 Security concern

There are potential security concerns that could be present in The Cage. The Cage has not been audited which means there could be underlying security concerns that go unnoticed, decreasing customer trust. One concern regarding security is data confidentiality. Customers may not be convinced that The Cage's hardware or database is safe enough to prevent others from accessing their personal data. Another concern is that customers may not be able to trust that The Cage will not share with a third party—a promise they make (Perdana et al., 2020). Moreover, The Cage relies on ERP systems and Cloud Storage, both of which have been vulnerable to hacking. As customers increase, the more data is generated and stored in the ERP system and cloud, thus putting greater amounts of data at risk of a security breach (Perdana et al., 2020 & White & Ariyachandra., 2016)

#### 3.1.3 Discrimination and monitoring issues

When it comes to using DA, the negligence of data governance is unethical since it impacts more than security and privacy issues but also discrimination and monitoring. Currently, The Cage only uses descriptive statistics to analyse data, but they want more useful data such as birth date and ethnicity of consumers to diversify their business (Perdana et al., 2020). However, collection of such data could be prone to discrimination. *Asadi Someh et al.* (2016) state that using DA to divide individuals into groups based on gender, ethnicity, colour, religion, and economic status, can result in giving or prohibiting priority privileges or services to a specific group, leading to discrimination. In regard to monitoring issues, The Cage's use of DA could result in monitoring issues, in the sense that customers could be tracked throughout multiple platforms (Asadi Someh et al., 2016).

### 3.2 To address the above problems, the recommendations are the following:

#### 3.2.1 Obey the PDPA regulations:

The Cage should certainly adhere to the PDPA because as they collect and analyze data their standards will be ethically consistent under PDPA regulation. The Cage should also obtain consent from the customers to perform certain actions on the data. Additionally, customers should be provided with the ability to withdraw their consent. Once consent is terminated, The Cage must stop collecting, analysing, and distributing a customer's personal data. Moreover, The Cage should be obligated to explain the collection and use of customer data to the customer, which cannot go beyond what is reasonable for the organisation to deliver a product or service.

### 3.2.2 Protection of data

When the corporation is processing personal data or disposing of documents containing personal data, The Cage should take appropriate security precautions to ensure that personal data protected while in their possession, including the storage devices on which such data is kept. According to *Howe & Elenberg (2020)*, Facebook's chief executive has announced actions the company is taking to improve customer privacy, one of which includes the adoption of encryption methods to safeguard what users see and send. This is done to prevent illegal access, collection, use, or revelation of such data.

### 3.2.3 Supervision of DA practices

In aid of the possibility of discrimination in DA, individuals in *The Cage* that handle customer data should be monitored to ensure the collection and use of data is discrimination and bias free. According to *Favaretto et al (2019)*, discrimination in DA can be mitigated through third-party supervision of DA practices. The Cage could hire a third party to supervise how they collect and use customer data.

### 3.2.4 Anonymization of DA

To mitigate privacy concerns that customers could hold about DA, The Cage could adopt anonymization methods in their DA. One method they could adopt is Privacy-Enhancing k-Anonymization Methods as referenced in *Zhong et al (2009)*. Privacy-Enhancing k-Anonymization allows individuals dealing with data to anonymize the data being handled (Zhong et al., 2009). In relation to The Cage, using Privacy-Enhancing k-Anonymization will allow customers of The Cage to ensure that their data collected is anonymous—protecting their privacy.

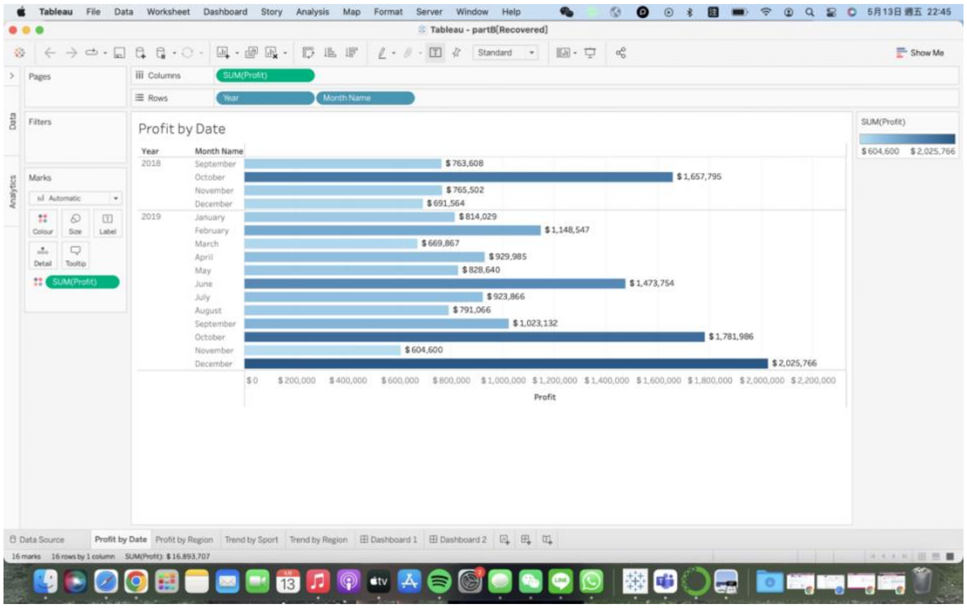
### 3.2.5 Enhance the awareness of ethics in applying data

According to *Asadi Someh et al. (2016)*, it is necessary to have a comprehensive understanding of ethical concerns and their consequences. Therefore, The Cage should hire a professional speaker or expert in the field of DA to deliver ethics training to their employees and affect their ethical consciousness.

## Part B

1. **Using Microsoft Power BI and the Excel file named 'MANG6526 Financial Sample', produce suitable charts/ graphs illustrating your response to the following questions:**
   a. **Which month and year had the most profit?**

According to Figure 1, The Cage made the most profit in October of 2018 ($1,657,795) and December of 2019 ($2,025,677). The most profitable year for The Cage was 2019. This could be attributed to how the data only includes four months of profit from 2018.

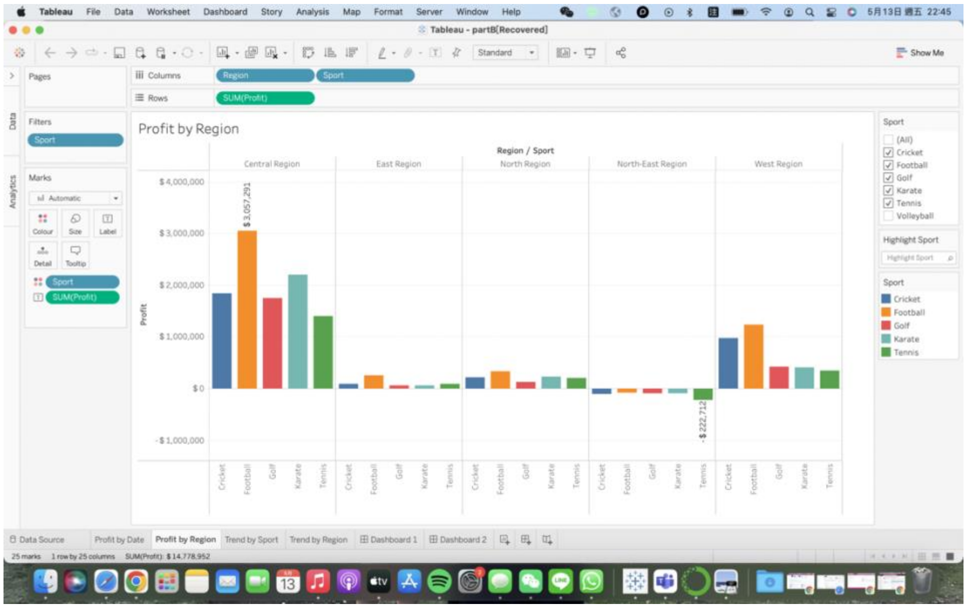



(Figure 1)

**b. The Cage discontinued providing 'volleyball' last month. Please filter this data from your report to avoid confusion. Which sport and region should the company continue to invest in? Which sports and regions, if any, should The Cage stop targeting/investing in?**

The Cage should continue to invest in Football in the Central Region as they had the highest profit at $3,057,291 (see Figure 2). The Cage should stop investing in Tennis in the North-east region because this sport in this region made a loss of $222,712 (see Figure 2).

(Tableau skills used: only show the label in the data selected)

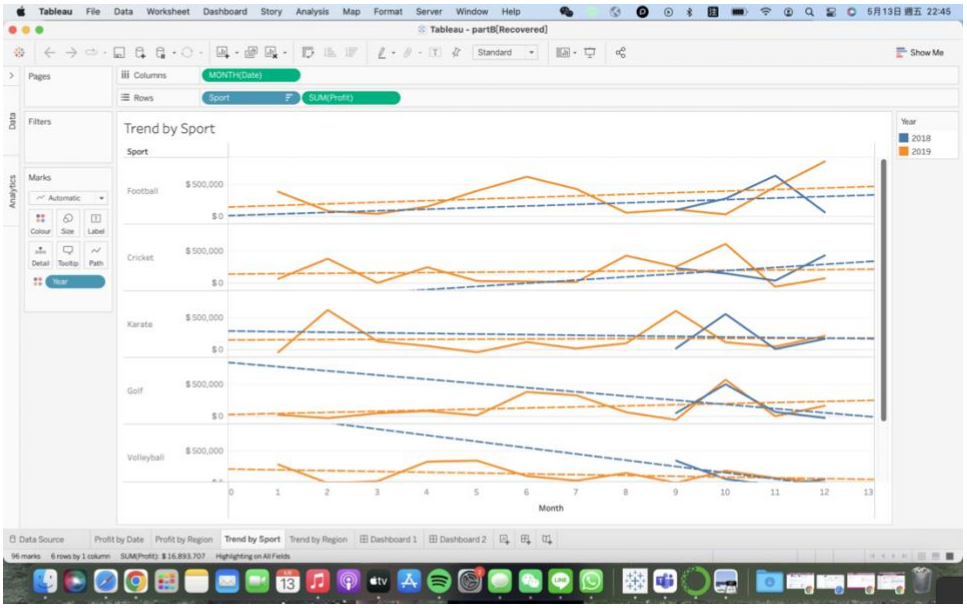



(Figure 2)

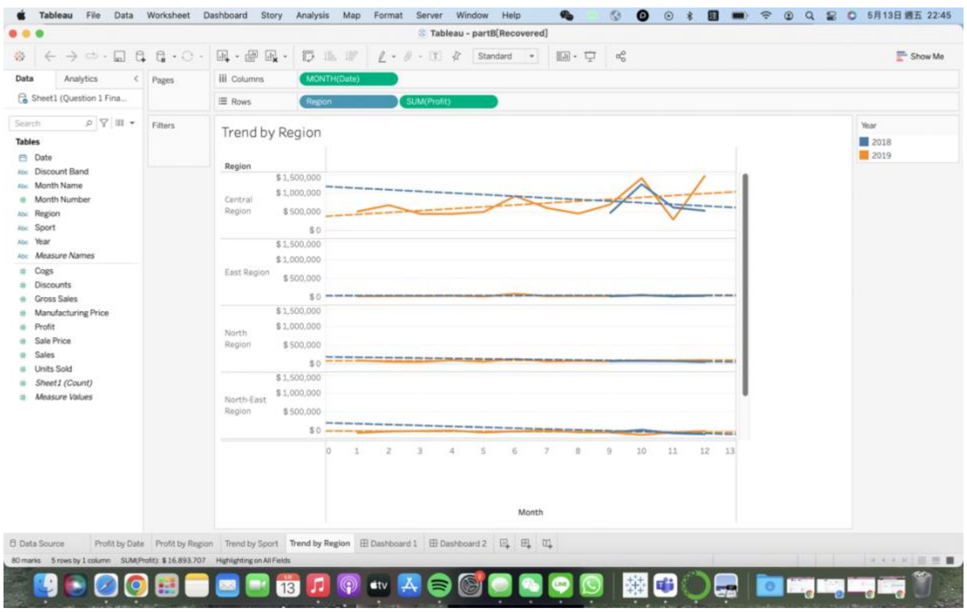
c. **Slicers are a valuable tool for filtering the visuals on a report to a specific selection. Using a Year Slicer, identify whether there have been any trend(s) over different Regions and sports throughout the past 15 months. Use screenshots to support your answer.**

According to Figure 3, Trend by Sport, the profits have risen for the consecutive months in two years for both Cricket and Tennis. Opposingly, Karate and Volleyball show a decreasing trend in both 2018 and 2019. Moreover, despite a decrease in golf in 2018, it has steadily increased in 2019. Among all sports, The Cage was most profitable (approximately $250,000-500,000) between October 2018 and 2019. This trend is expected to continue.

According to Figure 4, Trend by Region, the Central Region is on the rise in 2019 after a decline in 2018. This figure also displays more fluctuation in this Region, with a peak in October 2019. Profits in the East, North, North-East, and West regions have remained consistent, ranging from $0 to $500,000.




(Figure 3)

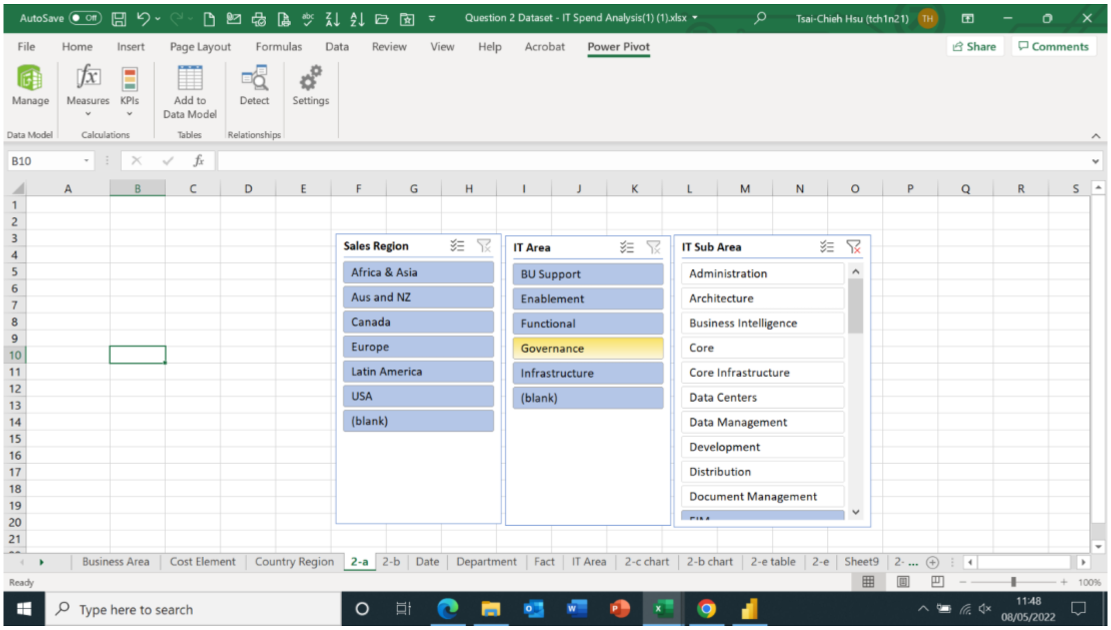



(Figure4)

2. **The Cage supplied you with their IT spend data. Using the available data, create a Microsoft Power BI report showing the following:**

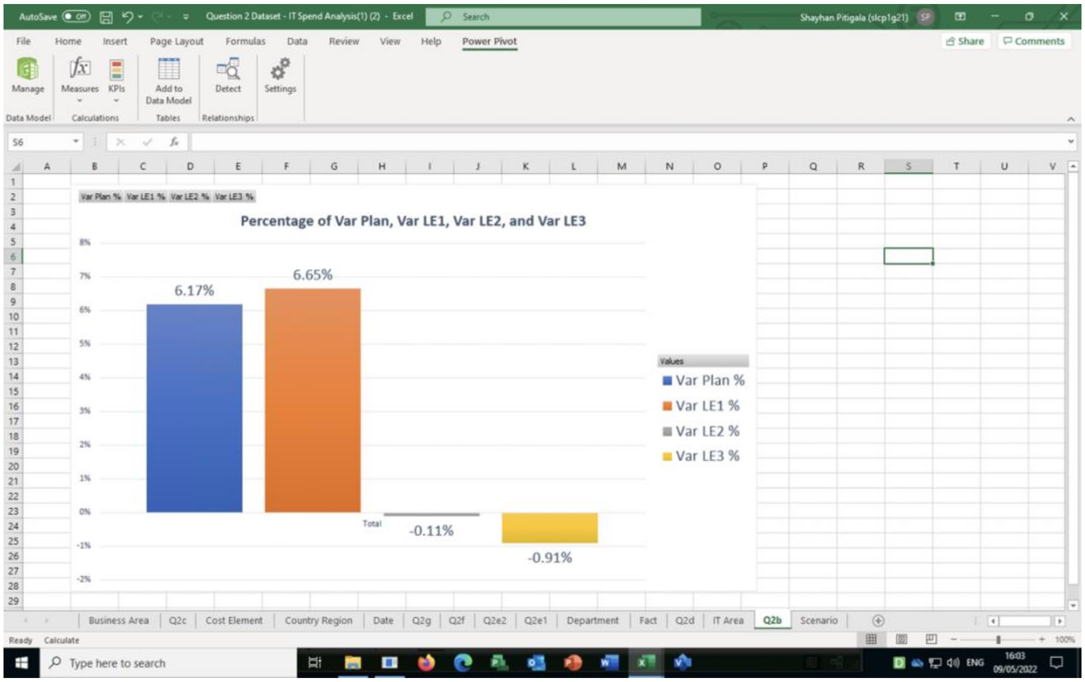
   a. **Add a slicer to filter data by the IT Area, IT Sub Area, and Sales Region data**

   Figure 5 refers to a slicer for the IT Area, IT Sub Area, and Sales Region data.


(Figure 5)

**b. Add a visual to represent the percentage of Var Plan, Var LE1, Var LE2, and Var LE3**

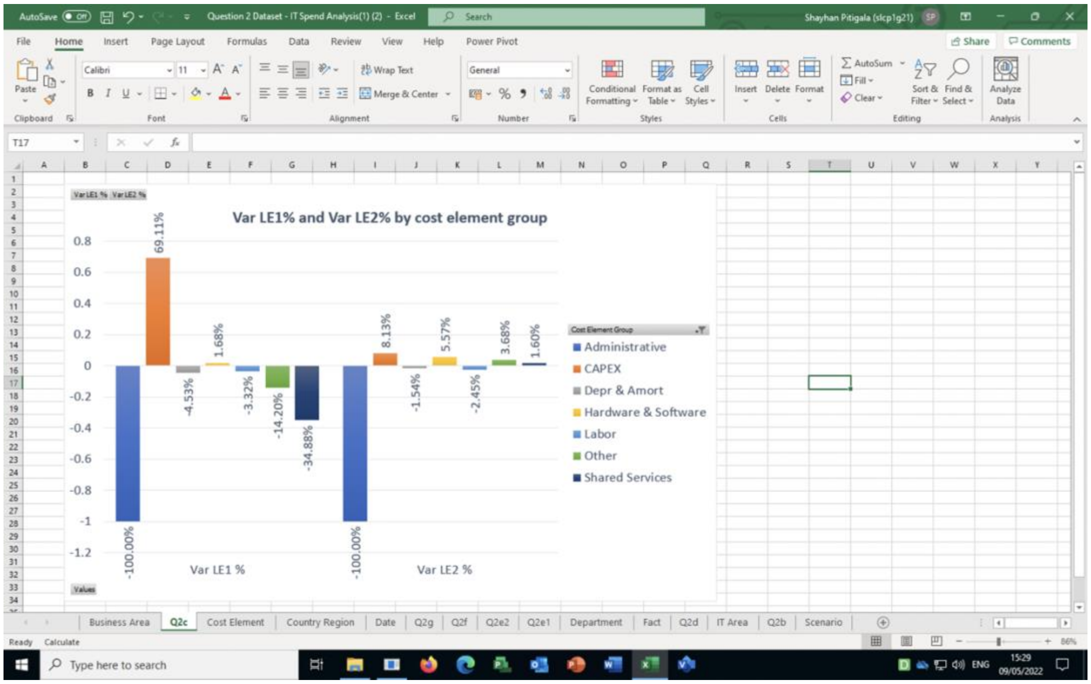
As seen in Figure 6, The Var LE1% is highest, meaning the actual spending is 6.65% greater that the estimated spending in LE1. The Var LE3% is lowest, meaning that the actual spending is 0.91% less that estimated in LE3.



(Figure 6)

c. **Add a visual to compare the values of Var LE1% and Var LE2% by cost element group**

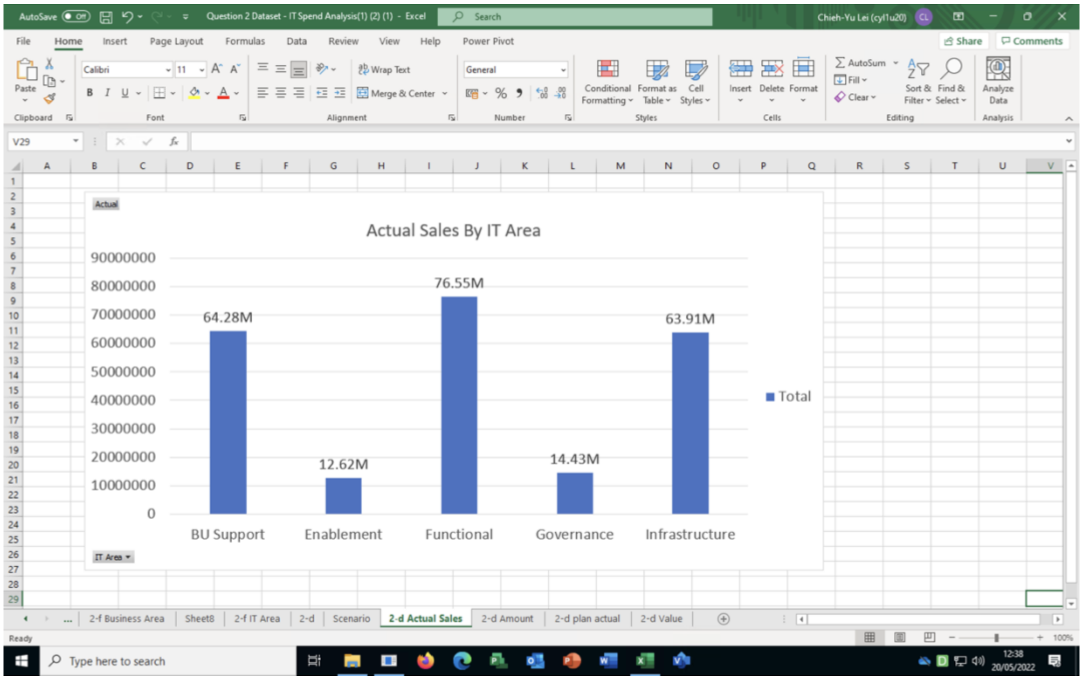
In reference to Figure 7, Among the cost element groups for Var LE1% and Var LE2%, the CAPEX cost element group cost most with 69.11% under the Actual in LE1 and 8.13% under the Actual in LE2. This means The Cage spent more than estimated on CAPEX. Administrative cost element group is negative 100% for both LE1% AND LE2%, meaning The Cage did not spend on this group.



(Figure 7)

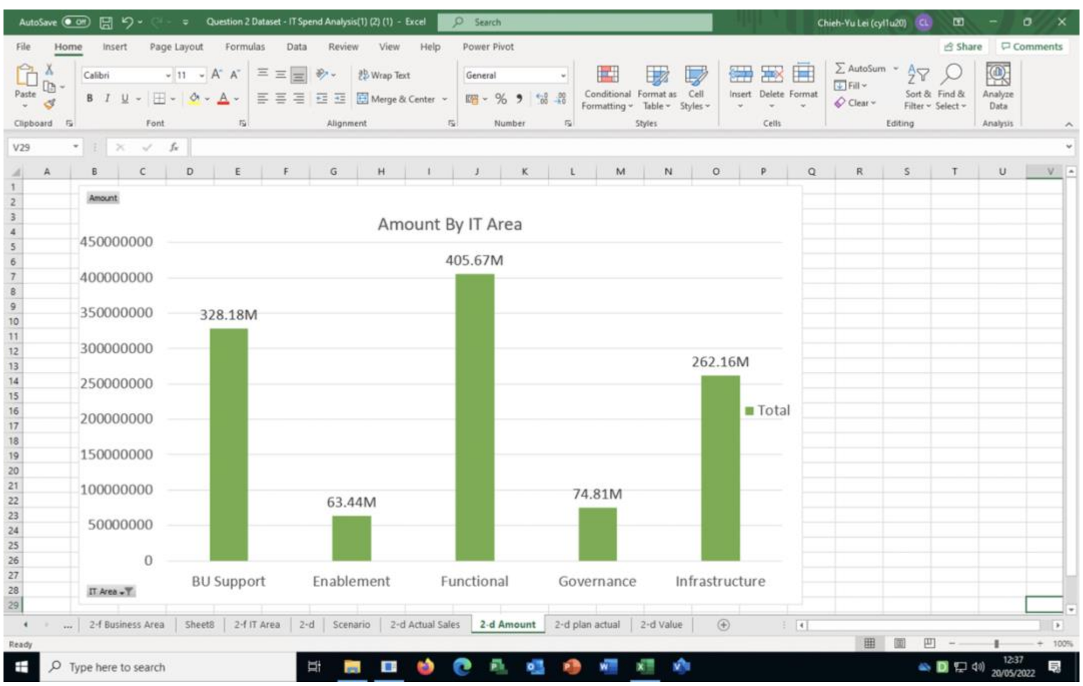
**d. Compare the Actual Sales, Actual/Plan, Amount, and Value information by IT area**

As seen in Figure 8, The Cage spent the most on the Functional IT Area at $76.55 million. They spent the least on Enablement and Governance at 12.62 million and 14.43 million. The company spent relatively similar amounts on BU Support and Infrastructure at 64.28 million and 63.91 million.



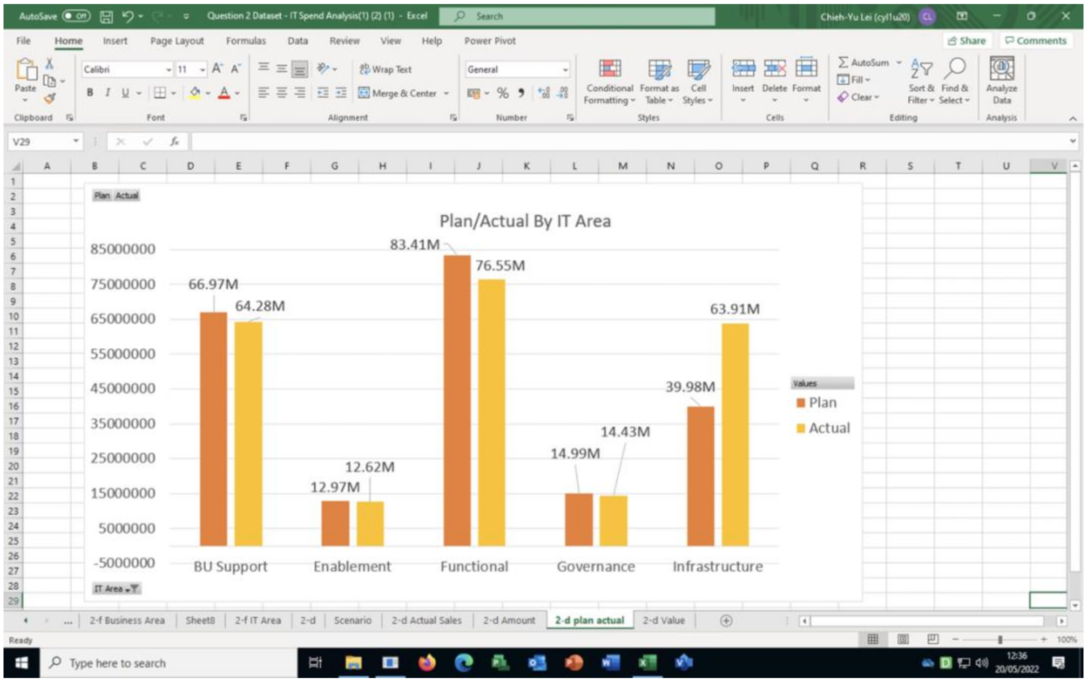
(Figure 8)

As seen in Figure 9, the IT Area with highest Amount in expenditure is the Functional IT Area at $405.67 million. The IT Area with lowest Amount in expenditure is the Enablement IT Area at $63.44 million.



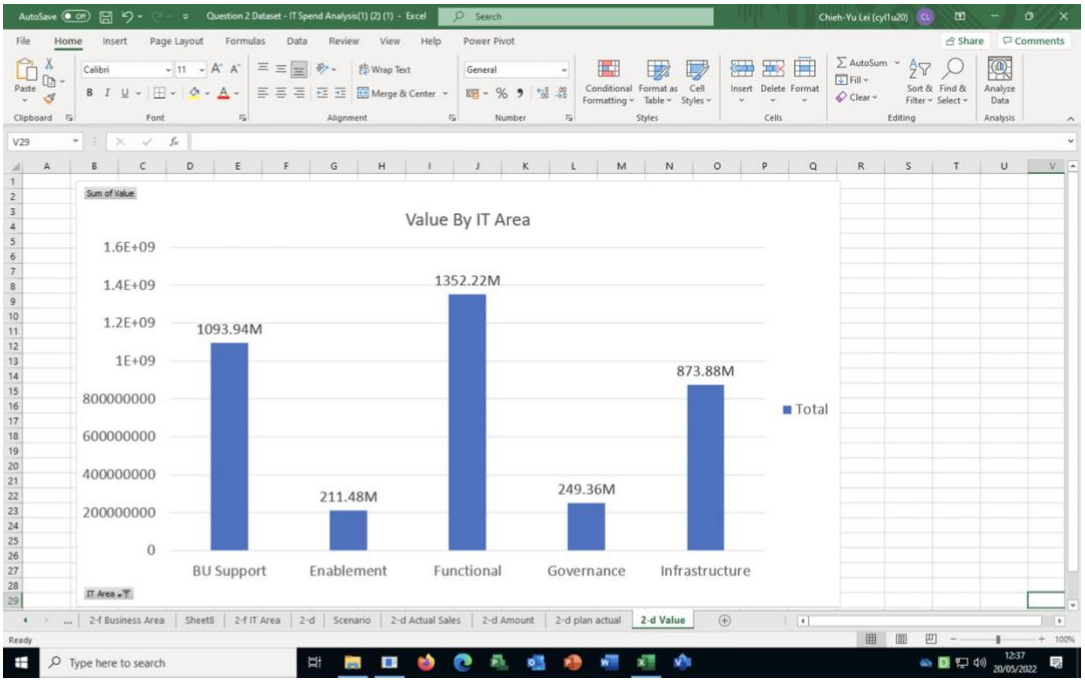
(Figure 9)

As seen in Figure 10, the spending between Actual and Plan scenarios are very similar among the categories: BU Support, Enablement, Functional, and Governance. However, in the Infrastructure IT Area, they spent $63.91 million in the Actual even though they spent $39.98 million in Plan, meaning they spent $23.93 million more in the actual than in the plan.



(Figure 10)

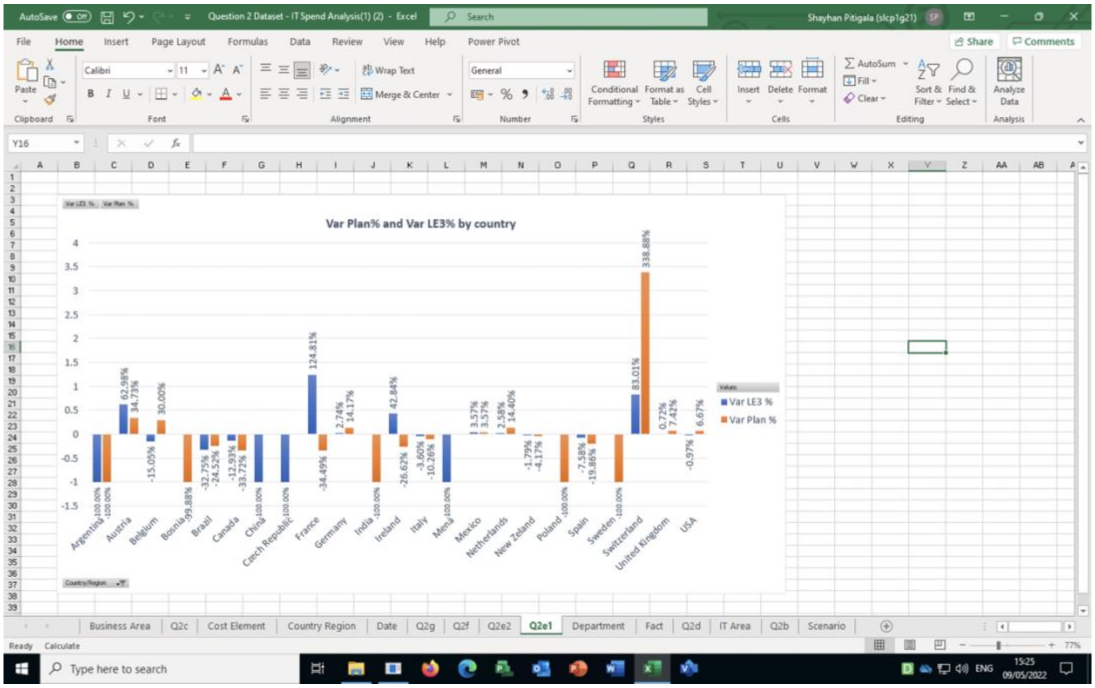
As seen in Figure 11, the Value is highest for the Functional IT Area at $1352.22 million, meaning that the company's Functional IT Area had the highest total expenditure out of all IT Areas. Opposingly, the Value is the lowest for the Enablement IT Area at 211.48 million, meaning that the Enablement IT Area had the lowest total expenditure out of all IT Areas.



(Figure 11)

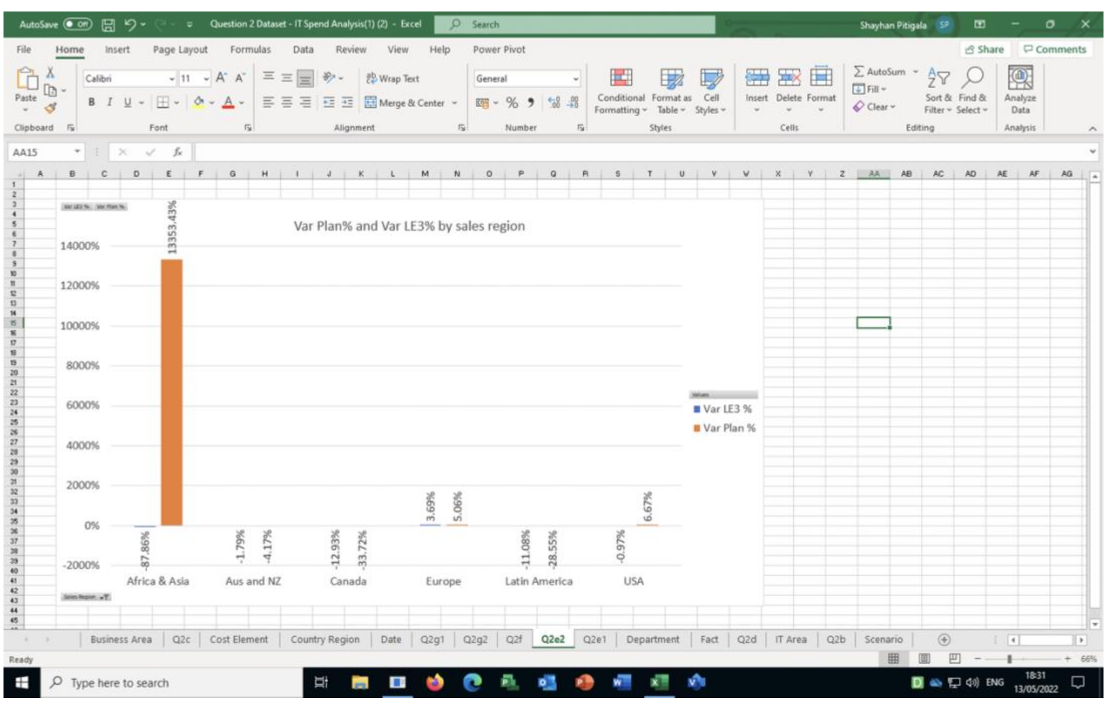
**e. Add a visual to represent the values of Var Plan% and Var LE3% by country and sales region**

As seen in Figure 12, Switzerland and France have the highest estimates with a Var Plan% of 338.88% for Switzerland and an LE3% of 124.81% for France indicating that the company underestimated their expenditure for these countries in these scenarios. Some countries are missing the LE3% and Plan% meaning the company did not estimate the expenses for these countries for these scenarios.




(Figure 12)

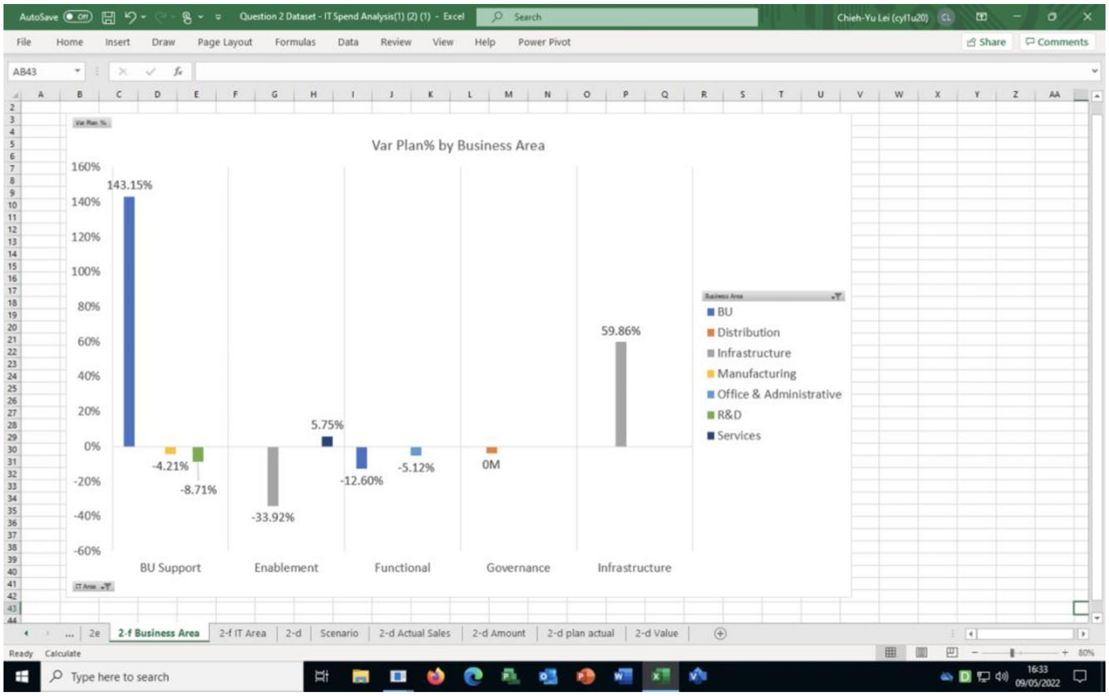
As seen in Figure 13, The Var Plan% for Africa & Asia is highest 13352.43%, meaning The Cage greatly underestimated their expenses for this sales region relative to the Plan scenario. Aside from Africa & Asia, in majority of sales regions, The Cage spent less in Actuality compared to the Var LE3 and Plan.




(Figure 13)

**f. Add a visual to compare the values of Var Plan% by IT area and business area**

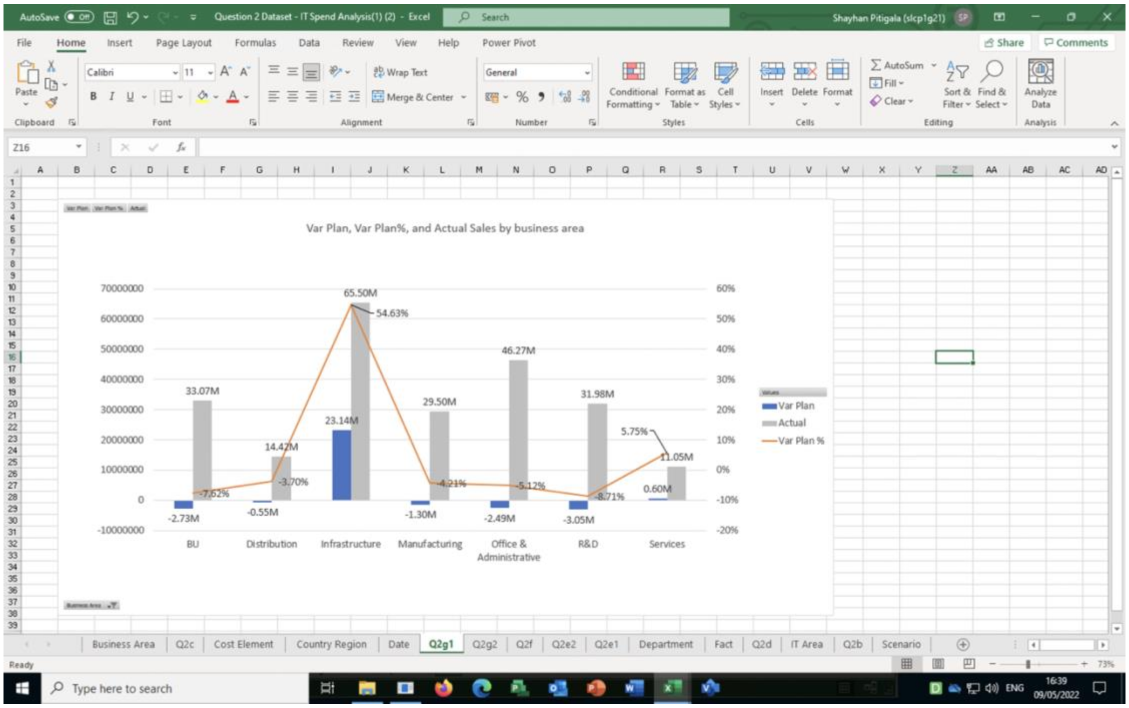
As seen in Figure 14, the Var Plan% is highest in the BU Support Area (including BU, Manufacturing, R&D) at 130.23% (143.15%-4.21%-8.71%), meaning that The Cage spent 130.23% more in the Actual than they spent in the Plan. This is due to the BU Support area containing the Business area "BU" which had a value of 143.15%, indicating that The Cage spent 143.15% more on BU in the Actual compared to the Plan. The Var Plan% is lowest for Enablement IT Area at -28.17% (-33.92%+5.75%) and the Infrastructure Business Area at -33.92%, meaning The Cage spent less for this IT Area and Business Area in the Actual than in the Plan.




(Figure 14)

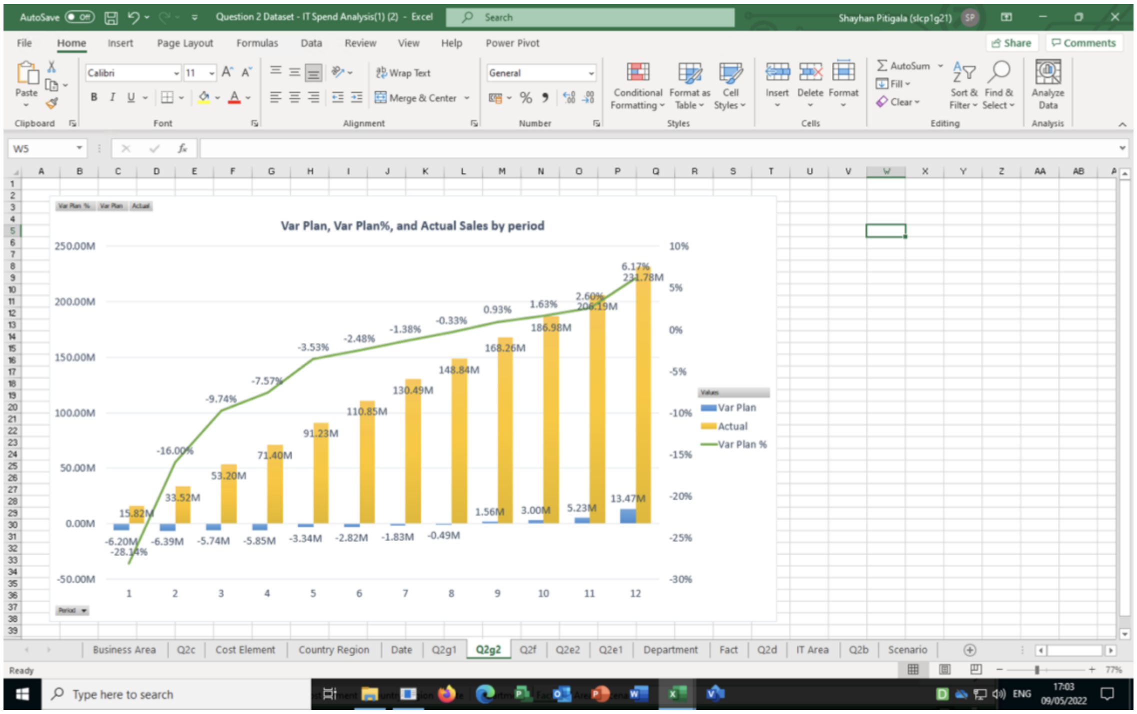
**g. Add a visual to present the Var Plan, Var Plan%, and Actual Sales by business area and period**

As seen in Figure 15, the Actual value is higher than the Var Plan for all Business Areas, meaning that the company has spent more than they estimated in the Var Plan Scenario. The Var Plan% peaks at 54.63% showing the largest difference between the actual and estimate spending in the Var Plan Scenario.



(Figure 15)

As seen in Figure 16, for every period (month), the Actual remains higher than the Var Plan, indicating that the company has spent more than they estimated over the course of 12 months. From months 7 to 10, the Var Plan% is +/- 2% indicating that the estimated spending is closer to the actual spending.



(Figure 16)

3. **The Cage supplied you with Budget, Forecast, and Actual data. Using the data provided, create a Microsoft Power BI report showing the IT spend over all areas of the business by doing the following:**

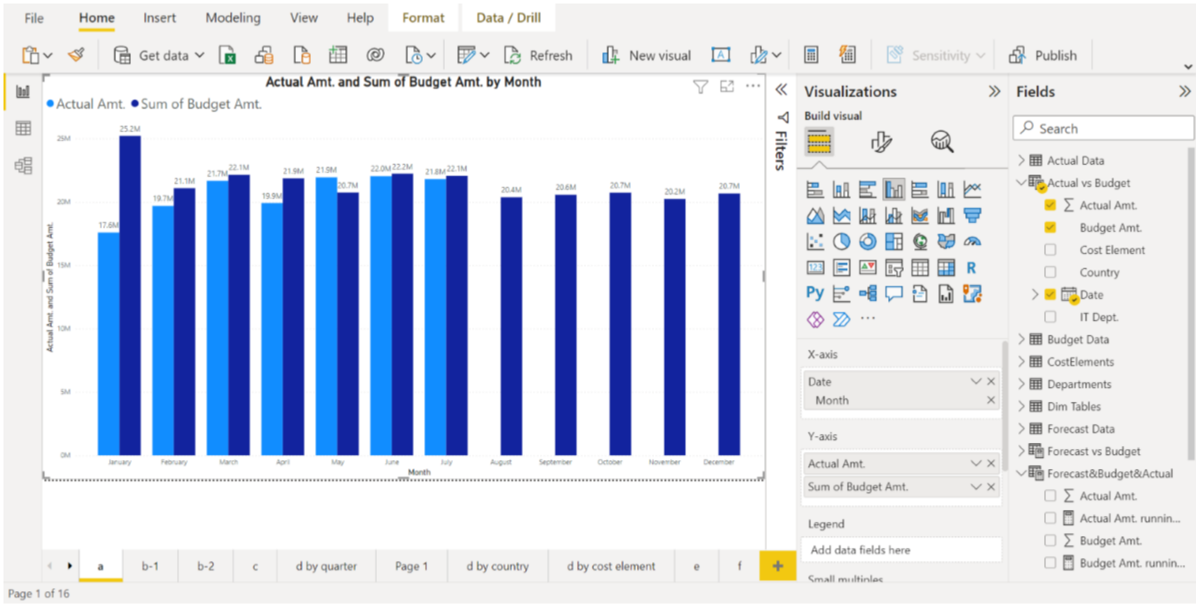
To compare the budget, forecast, and actual data by referring to dates, two steps are adopted to avoid many-to-many relationships and to link information properly. Firstly, as the actual data is separated in different files on a monthly basis, instead of importing files individually, we imported them by choosing the folder as our data resource. Consequently, new tables are conducted with DAX measures “UNION” and “SELECTCOLUMNS” to input and display two datasets in the same sheet. In the following tasks, these steps are pre-processed before demonstrating graphs.

a. **Add a visual to represent the Actual year to date versus the full year budget**

Figure 18 illustrates the comparison of actual and budget expenditures by month. Overall, the trend of the actual expense and budget are stable at around $20M (million). However, there is a larger discrepancy in January. We can tell that this data is potentially published in August as the actual expense is missing after then. Another thing we believe is worth noting is that the actual expense should be constrained by the budget given in the Cage, however, the actual expense exceeded the budget expense in May.

```
Actual vs Budget = UNION( SELECTCOLUMNS('Actual Data',"Date",'Actual Data'[Date],"IT Dept.",'Actual Data'[IT Department],"Cost Element",'Actual Data'[Cost Element],"Country",'Actual Data'[Country],"Actual Amt.",'Actual Data'[Actual],"Budget Amt.",BLANK()), SELECTCOLUMNS('Budget Data',"Date",'Budget Data'[Date],"IT Dept.",'Budget Data'[IT Dep.],"Cost Element",'Budget Data'[CostElement],"Country",'Budget Data'[Country],"Actual Amt.",BLANK(),"Budget Amt.",'Budget Data'[Budget]))
```

(Figure 17)



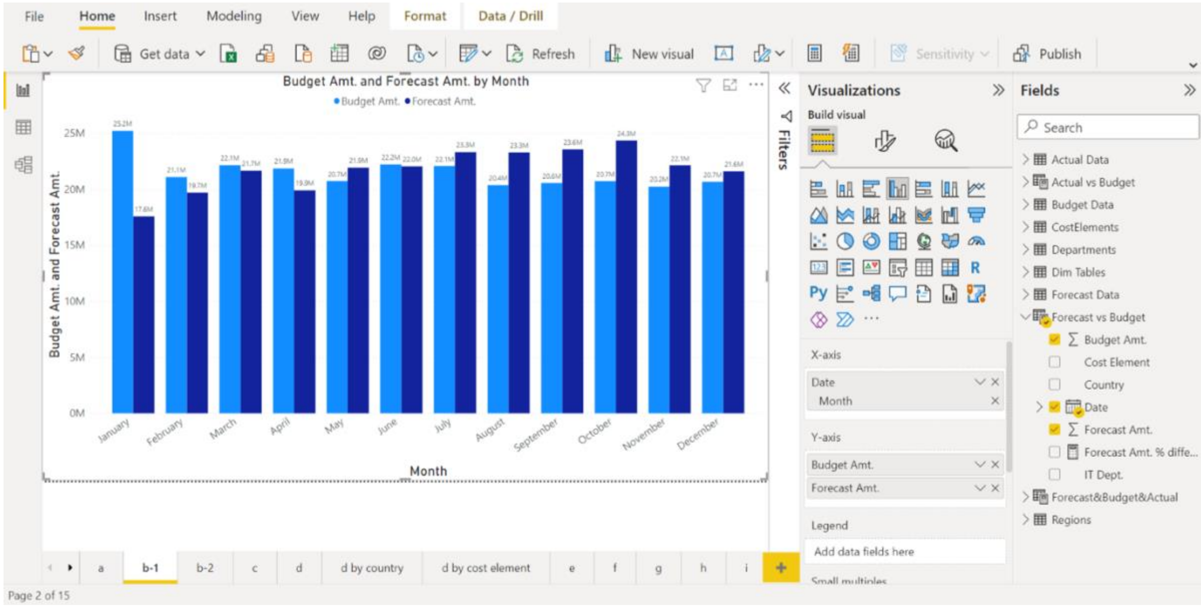
(Figure 18)

## b. Add a visual to represent the Forecast and Budget by data

As seen in Figure 20, gaps between the budget and forecast expenses are showing different trends. In the first four months, the forecast expense is lower than the budget expense. From then on, the forecast expense exceeds the budget expense. Since The Cage only referred to the historical data, they might ignore the budget amount in planning. Based on this, we recommend that The Cage should either decrease the forecast expenses or raise the budget to align with the target IT implementation plan.

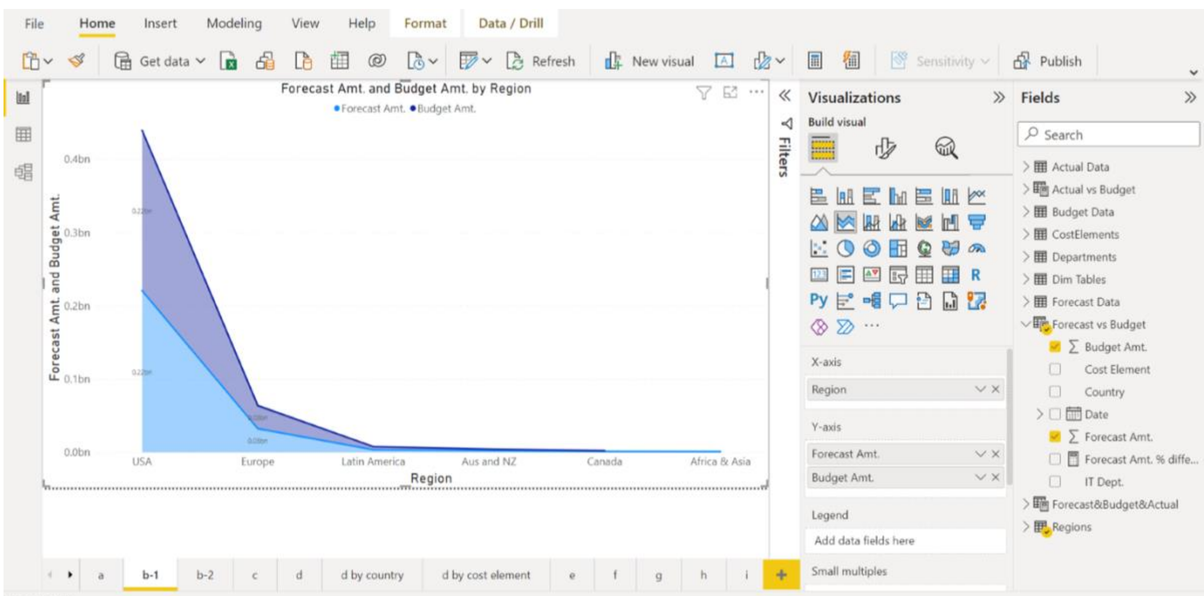
```
Forecast vs Budget = UNION (SELECTCOLUMNS('Forecast Data',"Date",'Forecast Data'[Date],"IT Dept.",'Forecast Data'[IT Dep.],"Cost Element",'Forecast Data'[CostElement],"Country",'Forecast Data'[Country],"Budget Amt.",BLANK(),"Forecast Amt.",'Forecast Data'[Forecast]), SELECTCOLUMNS('Budget Data',"Date",'Budget Data'[Date],"IT Dept.",'Budget Data'[IT Dep.],"Cost Element",'Budget Data'[CostElement],"Country",'Budget Data'[Country],"Budget Amt.",'Budget Data'[Budget],"Forecast Amt.",BLANK()))
```

(Figure 19)



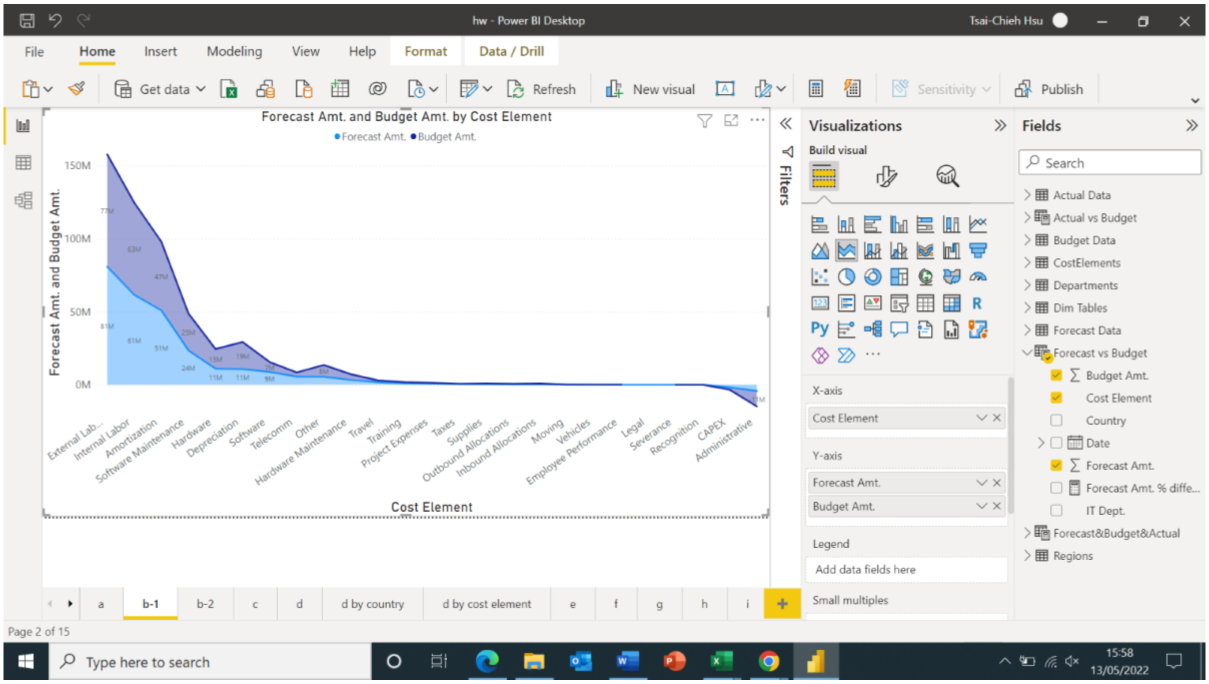
(Figure 20)

The forecast and budget for USA and Europe are similar as seen in Figure 21.



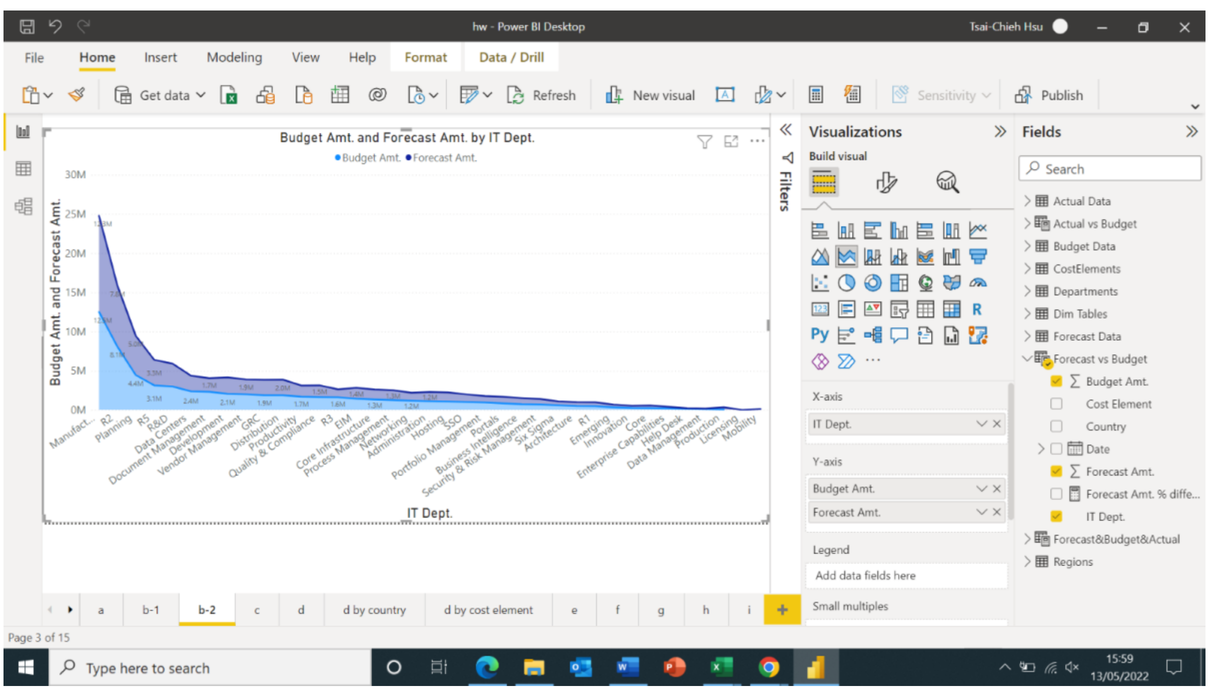
(Figure 21)

In Figure 22, the Cost elements with the highest Cost Elements (highest to lowest) include External Labor, Internal Labor, Amortization, and Software Maintenance.




(Figure 22)

According to Figure 23, the IT departments with the highest Budget and Forecast amounts include Manufacturing, R2, Planning, and RS. From then on, the Forecast and budget amounts follow a steady decline for all other IT Departments.



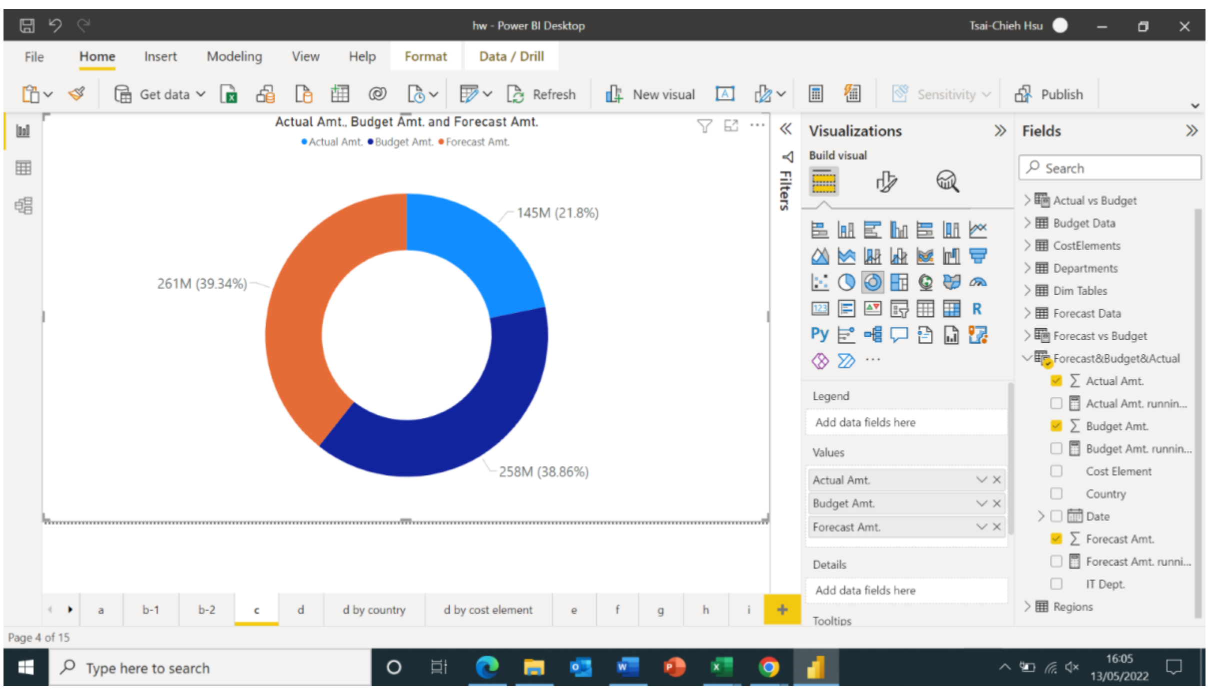

(Figure 23)

**c. Add a visual to represent the Actual, Budget, and Forecast by year, quarter, and month**

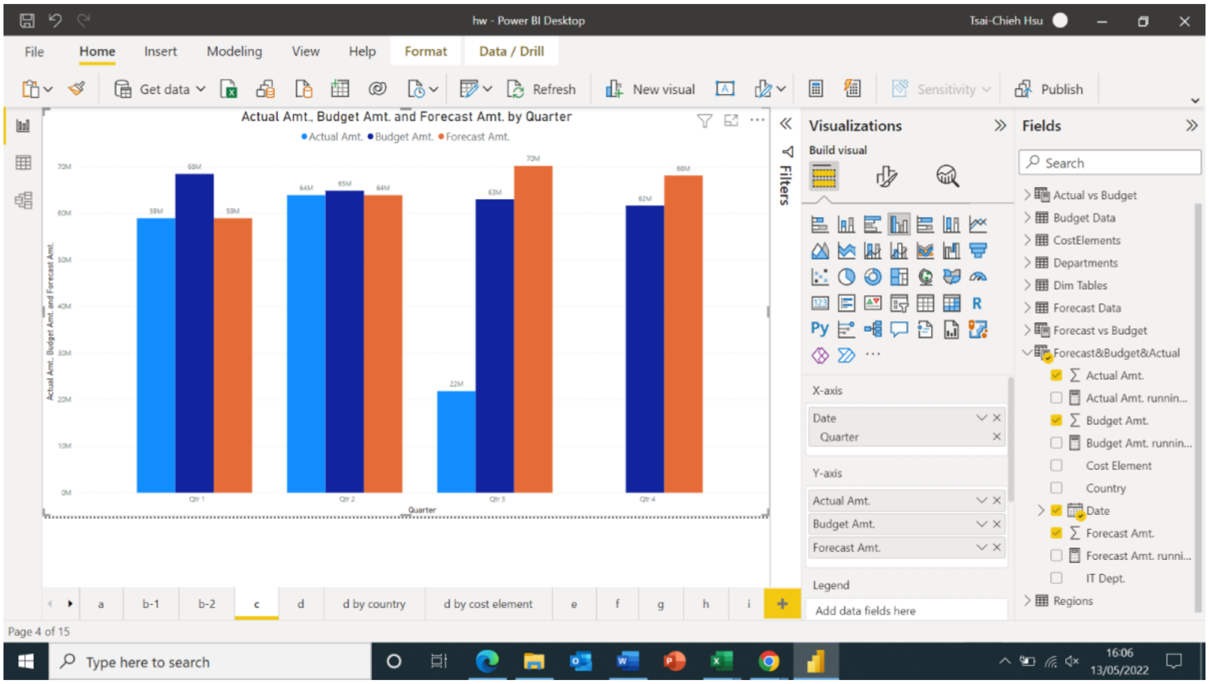
As seen in Figure 25, the total expenses of actual, budget, and forecast are 145M(21.8%), 258M(38.86%), and 261M(39.34%) respectively. As the actual expense only has seven months of data, we are not able to verify the reasonability that actual expense accounts for. Hence, the data is dissected into quarters and months. In the second quarter (See Figure 26), the differences between the three categories (Actual Budget, and Forecast) are slightly less than in the first quarter (See Figure 26). Looking at Figure 27, although the budget was higher in the first four months—especially for January which led to a large gap between actual and forecast expenses—the discrepancies in the other months are constant. We can imply that The Cage has improved their forecasting and budget control abilities.

```
Forecast&Budget&Actual = UNION(SELECTCOLUMNS('Actual Data',"Date",'Actual Data'[Date],"IT Dept.",
'Actual Data'[IT Department],"Cost Element",'Actual Data'[Cost Element],"Country",'Actual Data'
[Country],"Budget Amt.",BLANK(),"Forecast Amt.",BLANK(),"Actual Amt.",'Actual Data'[Actual]),
SELECTCOLUMNS('Forecast Data',"Date",'Forecast Data'[Date],"IT Dept.",'Forecast Data'[IT Dep.],
"Cost Element",'Forecast Data'[CostElement],"Country",'Forecast Data'[Country],"Budget Amt.",BLANK()
,"Forecast Amt.",'Forecast Data'[Forecast],"Actual Amt.",BLANK()),SELECTCOLUMNS('Budget Data',
"Date",'Budget Data'[Date],"IT Dept.",'Budget Data'[IT Dep.],"Cost Element",'Budget Data'
[CostElement],"Country",'Budget Data'[Country],"Budget Amt.",'Budget Data'[Budget],"Forecast Amt.",
BLANK(),"Actual Amt.",BLANK()))
```

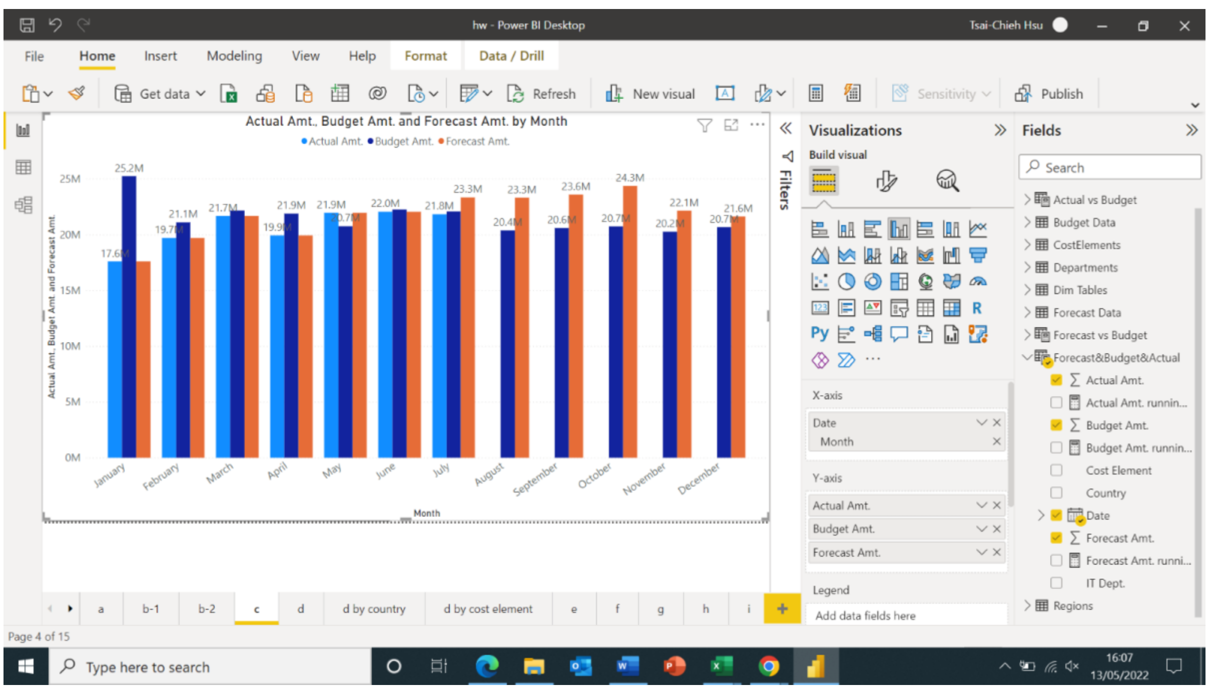
(Figure 24)



(Figure 25)




(Figure 26)




(Figure 27)

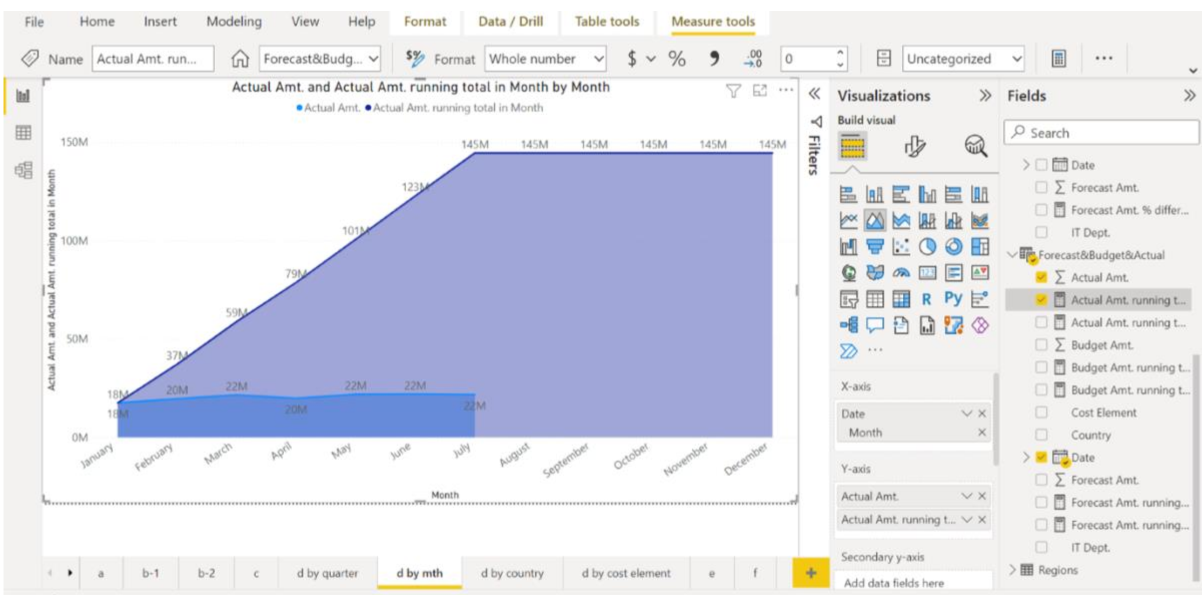
**d. Add a visual to represent the Running total for Actual, Budget and Forecast across dates – *you will need to write three DAX measures***

Quick measures aka “Running Total” were added using the calculation shown in Figure 28. Figure 29 shows that the actual expense grew at a linear increase in the first seven months until plateauing as a result of data not being developed at this point. As seen in

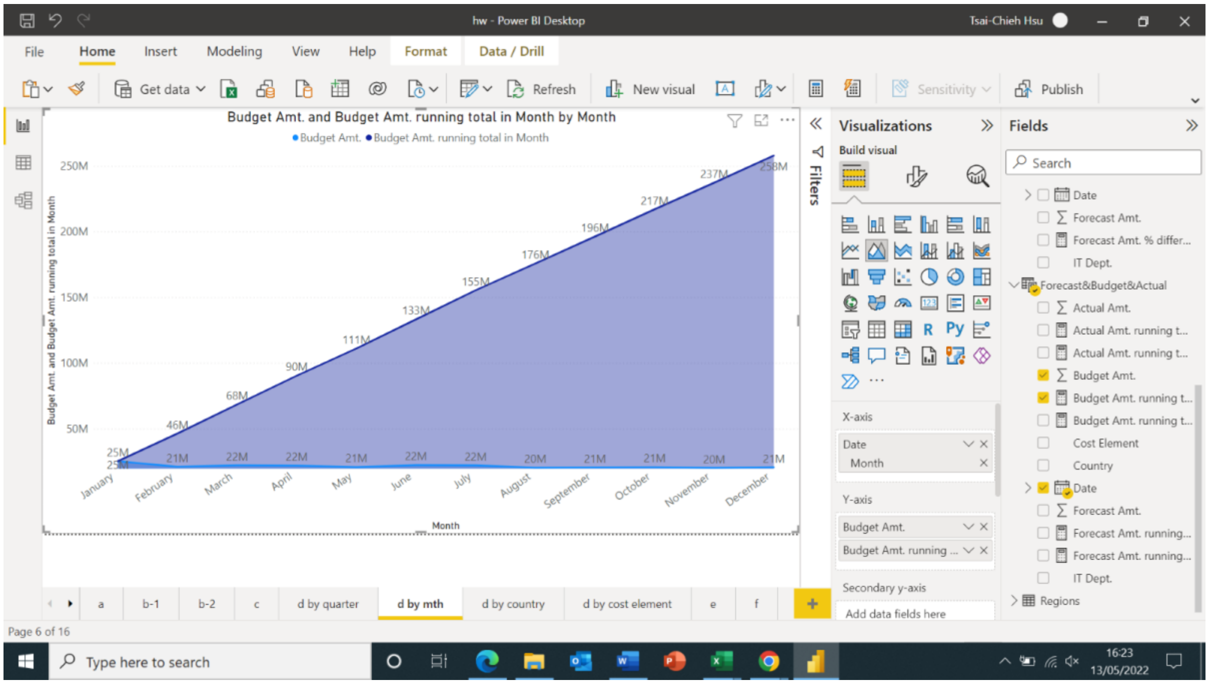
Figure 30, the forecast and budget expenses also show a linear increase with the expected annual expense at 261 and 258 million (seen in Figure 31). Referencing Figure 27 previously, As seen in Figure 32, comparing the three elements together (Actual, Forecast, and Budget Amount running totals by months), the Budget Amount expenses are slightly higher than the Actual and Forecast Amounts until October. We also see that even though the budget and forecast amounts increase, the actual amount plateaus after July. Meaning, that The Cage does not show its spending after July.

```
Actual Amt. running total in Month =
CALCULATE(
    SUM('Forecast&Budget&Actual'[Actual Amt.]),
    FILTER(
        CALCULATETABLE(
            SUMMARIZE(
                'Forecast&Budget&Actual',
                'Forecast&Budget&Actual'[Date].[MonthNo],
                'Forecast&Budget&Actual'[Date].[Month]
            ),
            ALLSELECTED('Forecast&Budget&Actual')
        ),
        ISONORAFTER(
            'Forecast&Budget&Actual'[Date].[MonthNo], MAX('Forecast&Budget&Actual'
[Date].[MonthNo]), DESC,
            'Forecast&Budget&Actual'[Date].[Month], MAX('Forecast&Budget&Actual'
[Date].[Month]), DESC
        )
    )
)
```

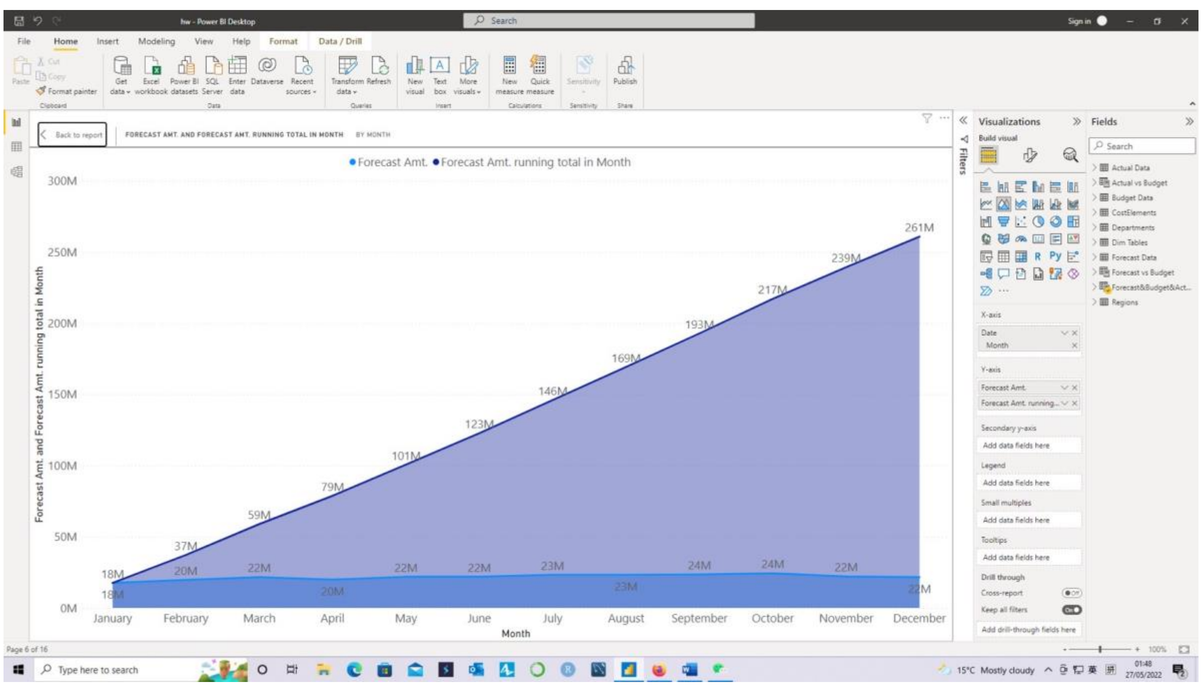
(Figure 28)



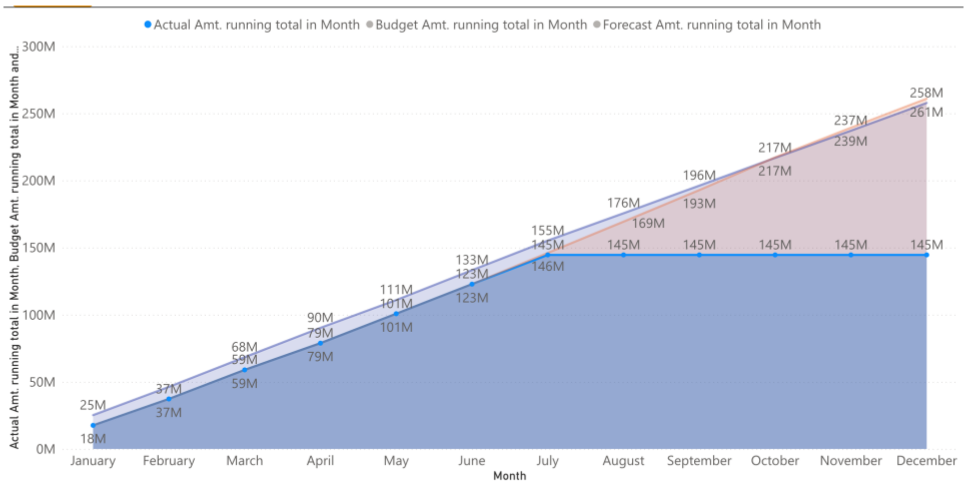

(Figure 29)




(Figure 30)

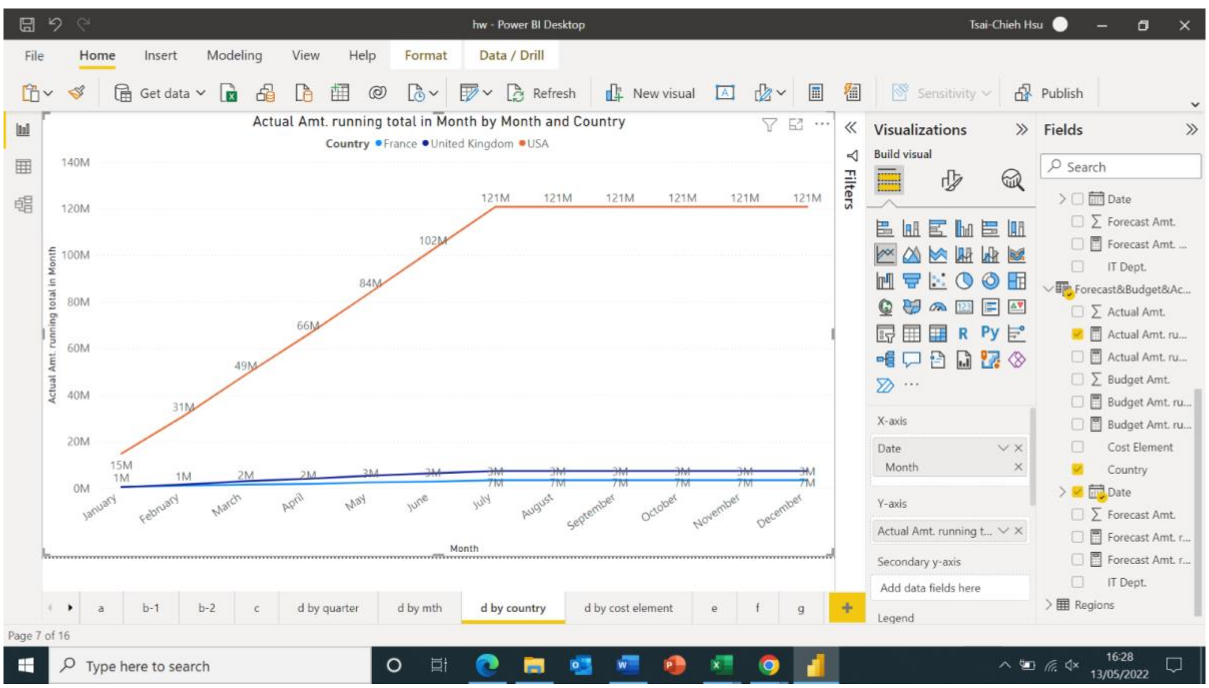



(Figure 31)



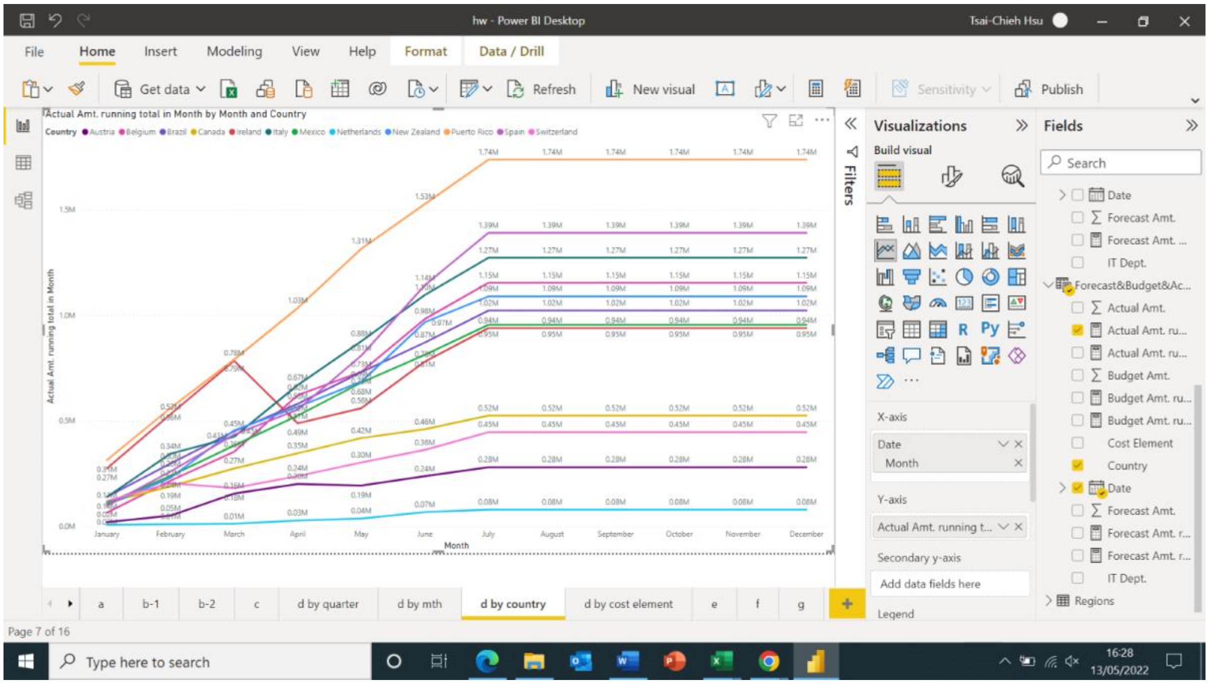

(Figure 32)

This graph use filter function to sort out the top three actual running total amount which are USA, United Kingdom, and France.



(Figure 33)

The rest of the countries are shown as below. One thing worth mentioning, the actual amount running total of Ireland downsized in April. This is because the actual amount in April is negative.



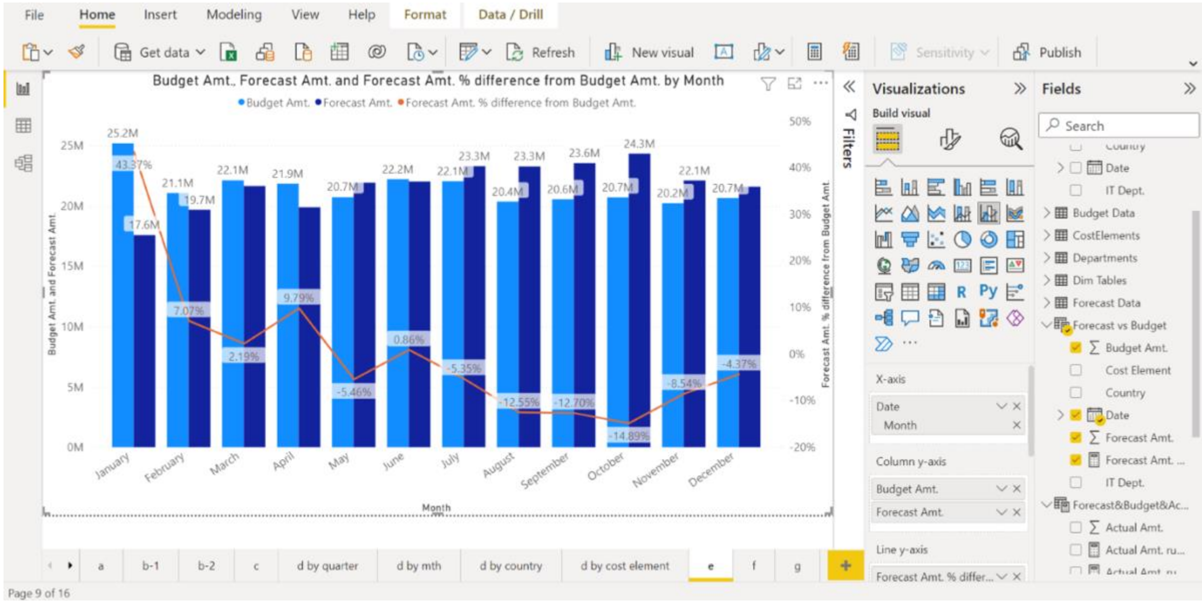

(Figure 34)

**e. Add a visual to represent the Budget versus Forecast and Budget versus Forecast % difference by date**

Figure 36 depicts budget and forecast expenses and their differences in percentage by month. A DAX measure was added to calculate the difference based on the forecast expenses (see Figure 35). For the most part, the Forecast Amount for The Cage follows a fluctuating decline from January to October.

```
Forecast Amt. % difference from Budget Amt. =
VAR __BASELINE_VALUE = SUM('Forecast vs Budget'[Forecast Amt.])
VAR __VALUE_TO_COMPARE = SUM('Forecast vs Budget'[Budget Amt.])
RETURN
    IF(
        NOT ISBLANK(__VALUE_TO_COMPARE),
        DIVIDE(__VALUE_TO_COMPARE - __BASELINE_VALUE, __BASELINE_VALUE)
    )
```

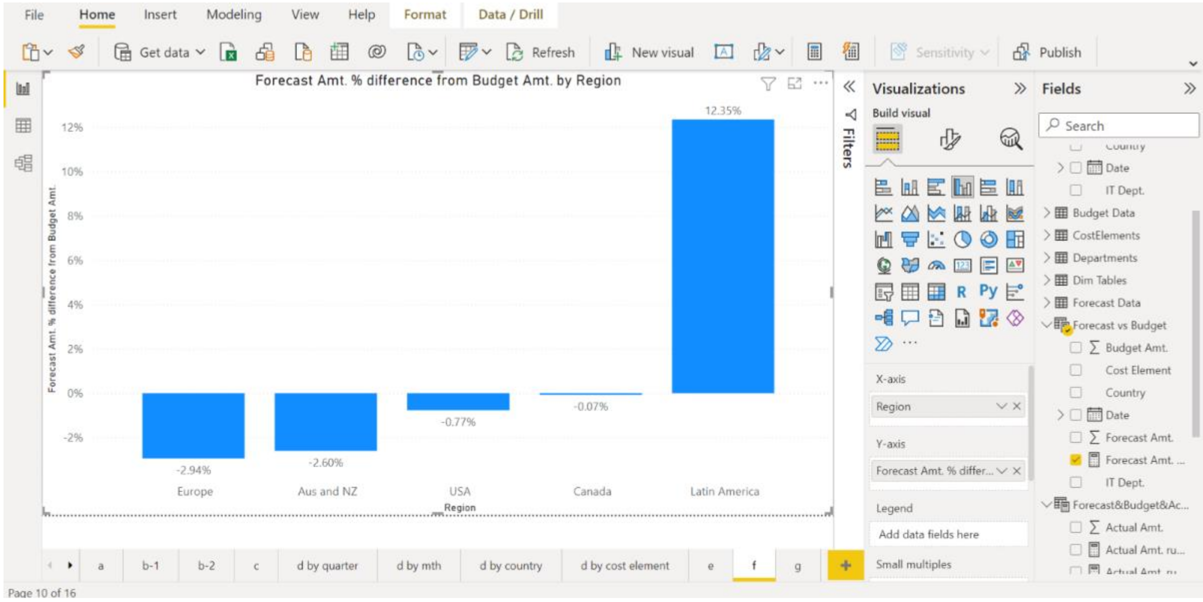
(Figure 35)



(Figure 36)

**f. Add a visual to represent the Budget versus Forecast % by Region, in ascending order**

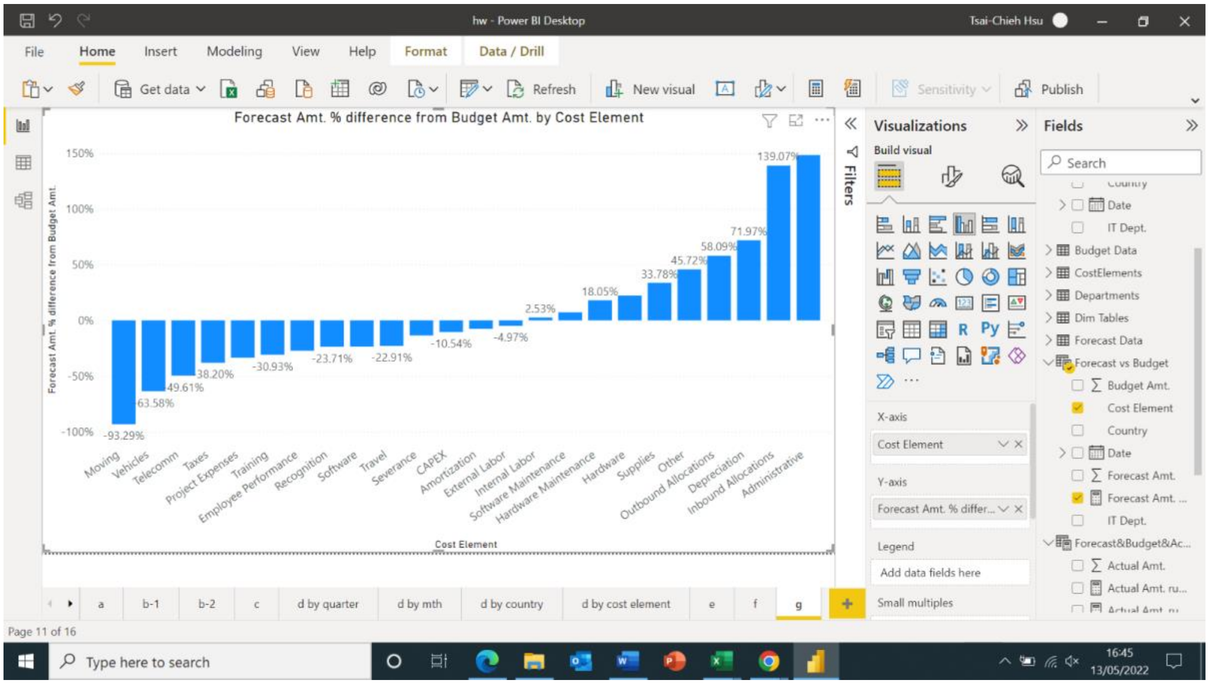
Figure 37 depicts the percentage difference between the Forecast amount from Budget amount in each region in ascending order. Apart from Latin America, the Forecast Amount % difference is negative among Europe, Australia and New Zealand, the United States, and Canada, meaning The Cage forecasted higher spending for the countries compared to the actual budget. Latin America has a Forecast Amount % difference from Budget Amount at 12.35%, meaning their budget was less than forecasted in this region. In contrast, Europe had the lowest Forecast amount % difference from Budget Amount at 2.94%, meaning their budget was higher than forecasted as should thus decrease the budget in this area.



(Figure 37)

**g. Add a visual to represent the Budget versus Forecast % by Cost Element, in ascending order**

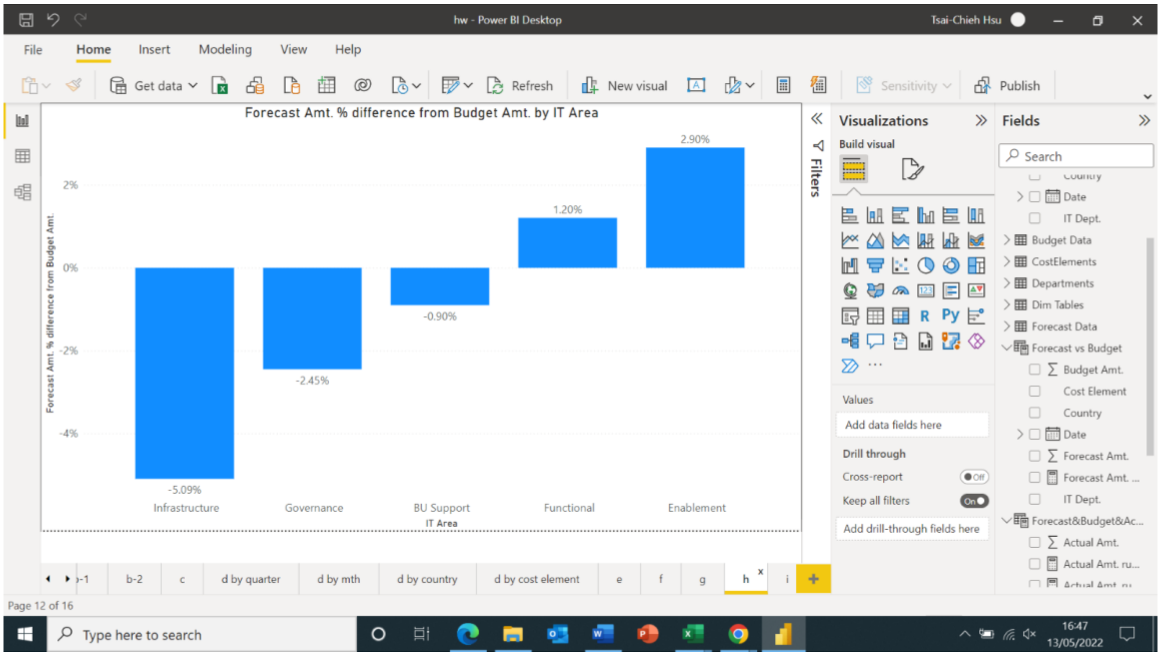
Figure 38 illustrates the Forecast Amount % difference percentage from Budget Amount by Cost Element. Out of all elements shown, the Moving cost element has the smallest % difference between the Forecast amount and Budget amount at -93.29%. This indicates that The Cage budgeted less for the Moving cost element than forecasted. Opposingly, The Administrative and Inbound Allocations have the largest % difference between the Forecast amount and Budget amount at 150% and 139.07%. This indicates that The Cage budgeted more for the administrative and Inbound Allocations than they forecasted.



(Figure 38)

**h. Add a visual to represent the Budget versus Forecast % by IT Area, in ascending order**

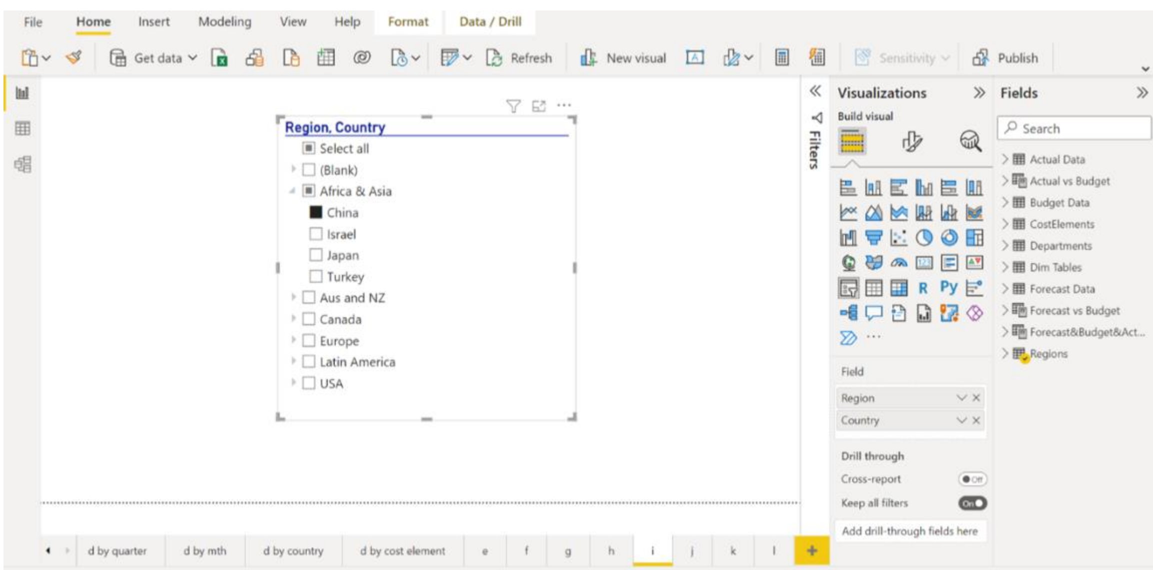
Figure 39 displays the % difference between Forecast amount and Budget amount by IT Area. The Infrastructure IT Area has the lowest % difference at –5.09%, indicating that The Cage budgeted less than forecasted by 5.09%. The Cage also budgeted less than forecasted for Governance and BU Support. However, BU support had a % difference of –0.90% which means the budget for BU Support was close to the forecasted amount. Opposingly, the Enablement IT Area had the highest % difference has 2.90%, indicating that The Cage budgeted more than forecasted for this IT Area. They also budgeted more for the Functional IT Area than forecasted.




(Figure 39)

## i. Add a slicer for the region

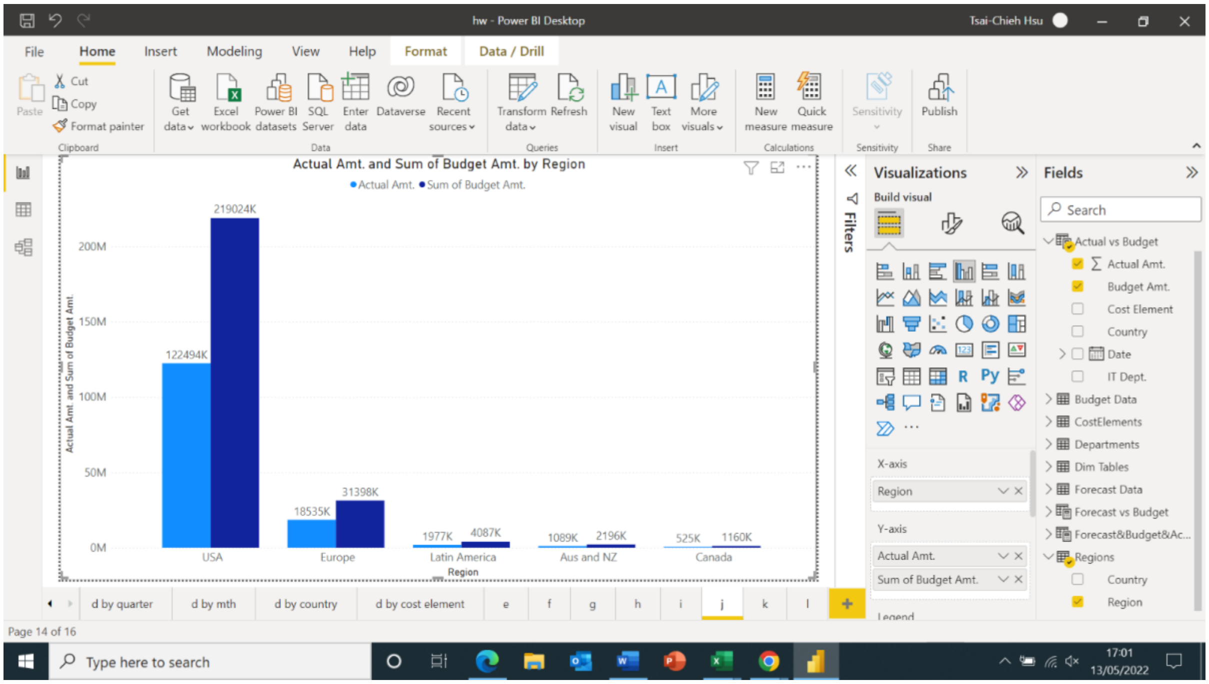
The slicer added here allows to select values between regions including Africa & Asia, Australia and New Zealand Canada, Europe, Latin America and the USA (see Figure 40).




(Figure 40)

## j. Add a visual to represent the Actual and budget by region

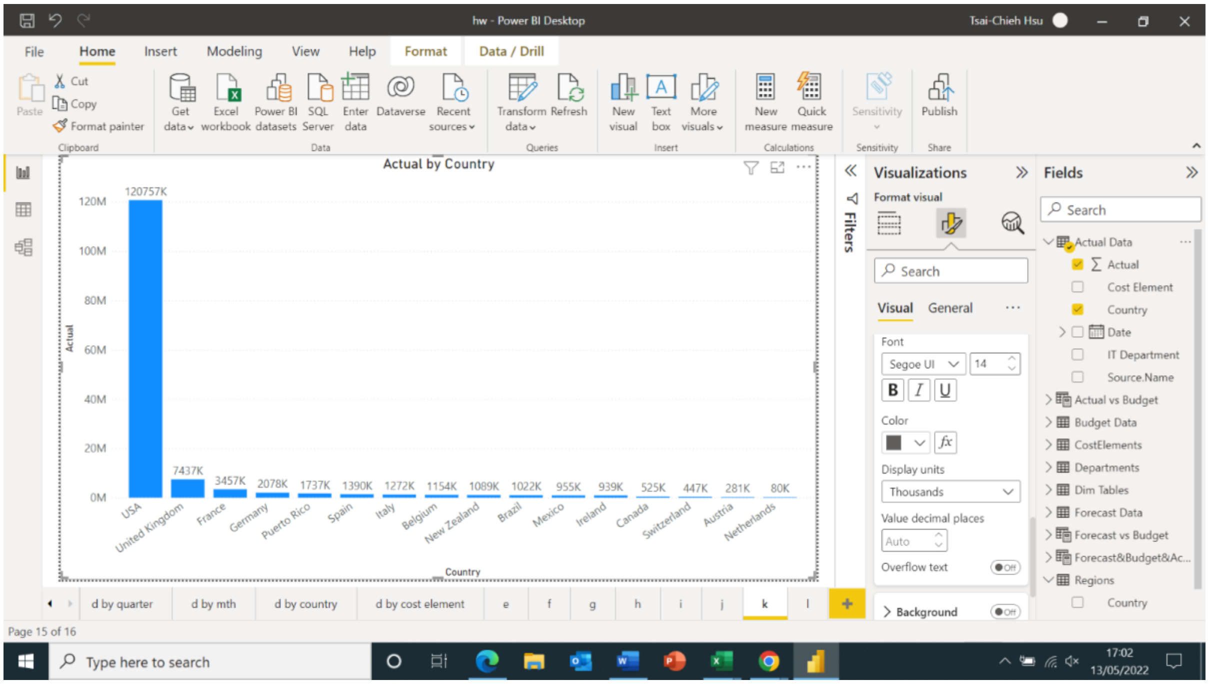
Figure 41 compares the actual Amount and Budget by Region. The USA had the highest budget at $219,024, but spent less with an actual amount of $122494, nearly twice as little as the budget for their region. Similarly, Europe only spent $18535 when their budget was $31400. As for Latin America, Australia and New Zealand, and Canada, they all had spent (Actual) without exceeding their budget.



(Figure 41)

**k. Add a visual to represent the Actual by country**

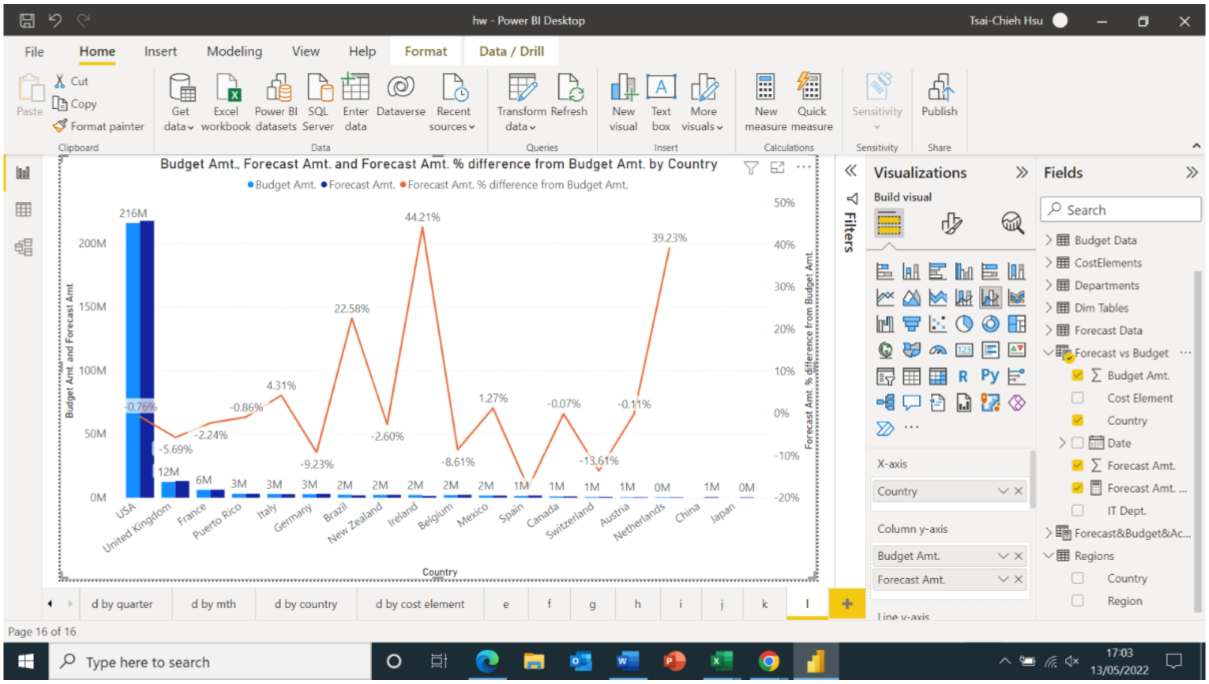
According to Figure 42, the USA spent the most in comparison to all other countries with an actual amount of $120757. That is $113,320 more than the United Kingdom in second and $120677 more than the Netherlands who spent the least.




(Figure 42)

## l. Add a visual to represent the Budget versus Forecast & Budget versus Forecast % by country

Figure 43 illustrates the budget compared to forecast and the budget compared to forecast % between different Countries. Nearly all budget from The Cage was allotted to USA at around $216 million, meaning that The Cage is prioritizing investment in IT in the United States. In addition, Ireland, Netherland, and Brazil have the highest % difference between Forecast and Budget (44.21%, 39.23% and 22.58%), meaning they all spent less than forecasted and have more room for IT investment. The Cage's allocation of budget is inconsistent among all countries. As a result, The Cage should distribute its budget evenly since they put a lot of focus on the USA.



(Figure 43)

Word Count: 6039

References

1. Adamopoulos, P., Ghose, A., & Todri, V. (2018) ‘The impact of user personality traits on word of mouth: Text-mining social media platforms’, *Information Systems Research*, 29(3), pp. 612–640. doi: 10.1287/isre.2017.0768.
2. Asadi Someh, I., Breidbach, C.F., Davern, M. and Shanks, G. (2016) 'Ethical implications of big data analytics’, *24th European Conference on Information Systems, ECIS 2016*.
3. August, T., Fox, R., Roy, D. and Pocock, M., (2020). ‘Data-derived metrics describing the behaviour of field-based citizen scientists provide insights for project design and modelling bias’, Scientific Reports, 10(1), pp.614-623. doi: 10.1038/s41598-020-67658-3.
4. Banerjee, A. and Basu, S., (2007). Topic Models over Text Streams: A Study of Batch and Online Unsupervised Learning. *Proceedings of the 2007 SIAM International Conference on Data Mining*. doi: 10.1137/1.9781611972771.40.
5. Buchanan, L. and O’Connell, A. (2006) ‘A Brief History of Decision Making’, *Harvard Business Review*, 84(1), pp. 32–41.
6. Casaló, L. V., Flavián, C. and Guinalíu, M. (2010) ‘Relationship quality, community promotion and brand loyalty in virtual communities: Evidence from free software communities’, *International Journal of Information Management*, 30(4), pp. 357–367. doi: 10.1016/j.ijinfomgt.2010.01.004.
7. Chan, H. K., Wang, X., Lacka, E., & Zhang, M. (2016) ‘A Mixed-Method Approach to Extracting the Value of Social Media Data’, *Production & Operations Management*, 25(3), pp. 568–583. doi: 10.1111/poms.12390.

8. Di Minin, E., Tenkanen, H. and Toivonen, T., (2015). ‘Prospects and challenges for social media data in conservation science’, *Frontiers in Environmental Science*, 3(SEP). doi: 10.3389/fenvs.2015.00063.
9. Drèze, X. and Zufryden, F. (2004) ‘Measurement of Online Visibility and Its Impact on Internet Traffic’, Journal of Interactive Marketing (John Wiley & Sons), 18(1), pp. 20–37. doi: 10.1002/dir.10072.
10. Epsilon.com. (2018). *New Epsilon research indicates 80% of consumers are more likely to make a purchase when brands offer personalized experiences*. [online] Available at: <https://www.epsilon.com/us/about-us/pressroom/new-epsilon-research-indicates-80-of-consumers-are-more-likely-to-make-a-purchase-when-brands-offer-personalized-experiences> [Accessed 26 May 2022].
11. Favaretto, M., De Clercq, E., and Elger, B.S. (2019) ‘Big Data and discrimination: perils, promises and solutions. A systematic review’, *Journal of Big Data*, 6(1), pp. 1–27. doi: 10.1186/s40537-019-0177-4.
12. Goswami, A., Bharathi, S. V., Raman, R., Kulkarni, A. V., Joseph, S., & Kelkar, B., (2013). Synergies between social media features and user engagement to enhance online brand visibility - A conceptual model. *International Journal of Engineering and Technology.* 5(3), pp.2705-2718.
13. Henard, D. H. and Szymanski, D. M. (2001) ‘Why Some New Products Are More Successful Than Others’, Journal of Marketing Research (JMR), 38(3), pp. 362–375. doi: 10.1509/jmkr.38.3.362.18861.
14. Howe III, E. G. and Elenberg, F. (2020) ‘Ethical Challenges Posed by Big Data’, *Innovations in Clinical Neuroscience*, 17(10–12), pp. 24–30.
15. Ilori, M. O. and Irefin, I. A. (1997) ‘Technology decision making in organisations’, *Technovation*, 17(3), p. 153. doi: 10.1016/S0166-4972(96)00086-7.
16. Jindal, K. and Aron, R. (2021) ‘A systematic study of sentiment analysis for social media data’, *Materials Today: Proceedings*. doi: 10.1016/j.matpr.2021.01.048.
17. Katila, R. and Ahuja, G. (2002) ‘Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction’, The Academy of Management Journal, 45(6), pp. 1183–1194.
18. Krishnan, V. and Ulrich, K., (2001). Product Development Decisions: A Review of the Literature. *Management Science*, 47(1), pp.1-21. doi: https://doi.org/10.1287/mnsc.47.1.1.10668
19. Labrecque, L. I., vor dem Esche, J., Mathwick, C., Novak, T. P., & Hofacker, C. F. (2013) ‘Consumer Power: Evolution in the Digital Age’, *Journal of Interactive Marketing*, 27(4), pp. 257–269. doi: 10.1016/j.intmar.2013.09.002.
20. Mäntylä, M. V., Graziotin, D. and Kuutila, M. (2018) ‘The evolution of sentiment analysis—A review of research topics, venues, and top cited papers’, *Computer Science Review*, 27, pp. 16–32. doi: 10.1016/j.cosrev.2017.10.002.

21. McKenna, B., Myers, M. D. and Newman, M. (2017) 'Social media in qualitative research: Challenges and recommendations', *Information and Organization*, 27(2), pp. 87–99. doi: 10.1016/j.infoandorg.2017.03.001.
22. Moe, W. W. and Schweidel, D. A. (2017) 'Opportunities for Innovation in Social Media Analytics', *Journal of Product Innovation Management*, 34(5), pp. 697–702. doi: 10.1111/jpim.12405.
23. Moise, D., Shestakov, D., Gudmundsson, G., & Amsaleg, L. (2013) 'Terabyte-scale image similarity search: Experience and best practice', *2013 IEEE International Conference on Big Data*, Big Data, 2013 IEEE International Conference on, pp. 674–682. doi: 10.1109/BigData.2013.6691637.
24. Mustafa, N. and Kingston, P. (2014) 'Organisational Decision-Making Behaviour: A Review of Decision-Making Theories', *Journal of Organisation & Human Behaviour*, 3(1), pp. 22–32. Available at: http://www.publishingindia.com/GetBrochure.aspx?query=UERGQnJvY2h1cmVzfC8yMTE0LnBkZnwvMjExNC5wZGY= (Accessed: 25 May 2022).
25. Patil, J., Wagh, G., Dhonde, P., & Dange, V. (2017). Aspect Base Sentimental Analysis For User Reviews Using Automatic Sentiment Analysis Techniques. *International Journal of Advance Engineering and Research Development (IJAERD)*, 4(10), 182–185. Available at http://www.ijaerd.com/index.php/IJAERD/article/view/3791 (Accessed: 27 May 2022)
26. Pearson, T. and Wegener, R., (2013). *Big Data: The organizational challenge*. [online] Bain & Company, Inc., p.1. Available at: <https://www.bain.com/contentassets/25c167a5149c42168994338f9dc99ffe/bain_brief_big_data_the_organizational_challenge.pdf> (Accessed 26 May 2022).
27. Piller, F. T. and Walcher, D. (2006) 'Toolkits for idea competitions: a novel method to integrate users in new product development', *R&D Management*, 36(3), pp. 307–318. doi: 10.1111/j.1467-9310.2006.00432.x.
28. Perdana, A., Gaffoor, J. and Lee, H. H. (2020) 'Getting data analytics on board at The Cage', *Journal of Information Technology Teaching Cases*, 10(1), pp. 29–34. doi: 10.1177/2043886919899407.
29. Schreck, T. and Keim, D. (2013) 'Visual Analysis of Social Media Data', Computer, 46(5), pp. 68–75. doi: 10.1109/MC.2012.430.
30. Supermetrics (2022) Supermetrics. Available at: https://supermetrics.com (Accessed: May 27, 2022).
31. This is the homepage of PDPC (no date). Available at: https://www.pdpc.gov.sg/ (Accessed: 25 May 2022).
32. Turpin, S. M. and Marais, M. A. (2004) 'Decision-making: Theory and practice', *Orion*, 20(2), pp. 143–159. doi: 10.5784/20-2-12.
33. Von Hippel, E. (1986) 'Lead Users: A Source of Novel Product Concepts', *Management Science*, 32(7), pp. 791–805. doi: https://doi.org/10.1287/mnsc.32.7.791

34. White, G. and Ariyachandra, T. (2016) 'Big Data and Ethics: Examining the Grey Areas of Big Data Analytics', *Issues in Information Systems*, 17(4), pp. 1–7.
35. Zhong, S., Yang, Z. and Chen, T. (2009) 'k-Anonymous data collection', *Information Sciences*, 179(17), pp. 2948–2963. doi: 10.1016/j.ins.2009.05.004.