Data Visualisation Feedback Form

|  |  |
| --- | --- |
| Report Title | What is the impact of the COVID-19 on the worldwide video streaming market? |
| Name/Id of student seeking | Hasan Guray (W1948912) |
| Name giving feedback | Yassine Kerkeni (W1930840) |
| Feedback given on (date) | 4 August 2023 |
| feedback Reviewer’s comments (what is good, what could be improved?) | |
| Research question | |
| The research question is well-elaborated and has the right context with the right aspect of the problem at hand. The research question spans the corresponding years that were put in question to study the case. | |
| Data acquisition | |
| The data source is reliable and has a good layout for checking the website and the relevance of the data. The acquisition process is well-identified and well-outlined in the report: Up to date data, with relevance, and the source is secure. Checking the source, I managed to identify that the data is accurate. | |
| Data preparation | |
| Power Query was used to identify data discrepancies and prepare the data.  Data profiling was used, and data cleaning: replacing NA values, transforming data types, and the use of the Power Query was used. The downsides of the preparation and transformation of the data are:   1. No advanced DAX was used to help identify new metrics; the different steps taken to clean the data do not present a level of complexity that matches the scope of the data.  The only measure used was the SUM function for different instances. 2. There were no calculated columns added to the dataset. | |
| Data analysis / Narrative | |
| * EDA: The EDA included some relevant steps: understanding the distribution of the data.  Understanding of the trends in the data. * The narrative is quite coherent, there are examples given and visualizations used to support the analysis.   However, there is a redundancy in visualizations that showcase the same thing [e.g.: Fig 9 and Fig 7 on the report] (this might be an unconscious mistake) but the analysis chunk preceding Fig.9. is talking about another metric. (This is to check again) | |
| Visualisation | |
| 15 visualizations were used, most of them respect the dashboarding design rules:   * 5-second rule to understand the visual. * User-friendly / Interactive use of the visuals on PowerBI (using slicers), use of a significantly appealing color palette.   However, some of the other visuals fail to respect these rules:   * Figure 11: The line graph is hard to understand as there are 12 services and the visual looks crowder than it should be. It is not easy to understand which line is for which service. * Figure 13: The map from PowerBI is a powerful tool, and this is a great visual that attracts the report reader, however, in the visual, there was no MAP KEY used which makes it hard for the reader to understand what each color signifies. Recommendation: add a key map to the visual. * Figure 14 & Figure 16: Represent a count of the countries that witnessed a growth rate by revenue and Penetration rate respectively: the use of the pie chart and the donut chart is not the best in these cases: counts could be presented in a simple clustered column chart that has both the metrics in the same visual. The use of these visuals is advised when there are more than 2 categories. | |
| Reviewee’s comments (What – if anything – did you change in your report after the feedback?) | |
| Before submitting coursework for evaluation, the practice of asking for feedback from our colleagues has proven incredibly beneficial. A detailed examination was then undertaken of their input; necessary cautions were then implemented in response to each identified concern raised.  Primarily, an exhaustive assessment was conducted concerning the facets that were found to be deficient in the realm of visualization. Notably, in the instance of Figure 11, a transition of 12 services to subscribers was observed from 2017 to 2022 (Figure 1). After it became evident that the diversity of services under examination impeded the ability to trace certain lines, a deliberate refinement has been implemented, narrowing the analysis focus down to five of them based on their 2022 subscriber volumes.  A graph with colored lines  Description automatically generated  **Figure 1** *This is a screenshot of the Subscribers Trend by Service line graph from Power BI.*  Figure 13 demonstrated an ineffective map representation due to its absence of a legend or map key, making its color scheme difficult to comprehend. Thus, an appropriate legend or map key has been included to remedy any informational gaps within this visual.  Pie and donut diagrams presented in Figures 14 and 16 were chosen because of their clarity and impactful presentation, aiding the simultaneous understanding of numerical values and proportions. Their selection was further informed by seeking variety among graphic representations; employing clustered column charts was avoided due to potential complications arising from multiple categorical classifications within one graph.  An inadvertent misappropriation was identified in the graphic presentation, where the Penetration Rate Growth chart erroneously substituted the intended Total Users Growth chart within Figure 9. This discrepancy was promptly corrected thanks to feedback intervention.  Notably, computational methodologies employed via Data Analysis Expressions (DAX) remained manageably simple due to an appropriate dataset for graph formulation and research query explanation. A closer inspection of Power Query procedures showed a variety of transformative steps taken such as adding custom columns, renaming, removing columns, changing type, promoting headers, and unpivoting columns; with such preprocessing steps done within Power Query itself becoming unnecessary for advanced techniques of measure utilization.  In short, the suggestions from my colleague's feedback helped improve my work. By going through this step-by-step process, I was able to find and fix mistakes before submitting the coursework. This resulted in a better and more successful outcome. | |
|  | |