Who is leading the mobile device market?

Pooja Vijay Mahajan
pmah0895@uni.sydney.edu.au

***Abstract*— Mobile devices have become an essential part of human's life and are great to interconnect with people across the globe. The mobile device market continues to expand as the companies try to come up with ground-breaking innovations adhering to the consumer's demands. Aligning with this notion, the objective of this project is to visualize the leading mobile-device market. This report begins with an introduction section to describe the dataset, planning, implementation, and evaluation of visualization diagrams. Next sections illustrate non-visualization related individual contributions and visualization related contributions followed by key concepts of this unit (COMP5048) used in this project.**

I. INTRODUCTION

The advent of mobile devices is a result of ongoing evolution in technology [1]. The companies are competing against one another to create a mobile device with new capabilities. Increased speed, bigger chips, smaller transistors are used to boost the performance of the device [2]. Improvised detailing in the form factors such as dimensions, pixel density, depth has added to the customized versions of the devices in the existing market. In this visual analytics project, Mobile-Device dataset is used which contains a list of mobile devices introduced in the market in the period 1989-2012 and their respective features. There are 12 attributes namely: RAM, CPU, Storage capacity, Display diagonal, Display width, Display length, Width, Length, Depth, Volume, Mass, Pixel density.

The main objective of this project is to address three main aspects: devices creating new markets, companies leading new markets and the most successful company leading the market. After brainstorming several design approaches, our group decided to use Tableau to derive the visualizations. After thorough analysis and evaluation, the target was met, and we concluded that Samsung is the most successful company (refer to Figure 13 in group report).

II. NON-VISUALIZATION RELATED INDIVIDUAL CONTRIBUTION

This section enlists the non-visualization tasks that I have contributed towards and how I implemented them. My substantial contribution was towards the following tasks: Project Management, Data Analysis, Data Pre-processing, Literature Survey and Document Management.

A. Project Management

Once the groups were finalized, I reached out to my team members individually to get their contact email ID and phone numbers. I took everyone's consent and created a common communication platform (GroupMe and Outlook) for subsequent discussions. Seeking everyone's opinion, I also created a Work-Breakdown-Structure for our project. To add to that, I created a timeline using Gantt Chart to keep track of the deliverables and status of the project (Figure 1).

B. Data Analysis

I analysed the different columns in the 3 sheets namely, Sheet 1: Normalized Product Data, Sheet 2: Model-Company, Sheet 3: Company ID. I documented my inferences as supporting material for data pre-processing. Sheet 1 had 3162 rows comprising of different models of the mobile devices whereas Sheet 2 had rows with Model Id ranging from 260-3162 and hence had to be combined with Sheet 1 based on a unique column Model Id. Sheet 3 had 213 rows in which Old Id and Company were unique. Some of the companies merged and it resulted into a column called Company_real which had their own unique ID under the column New_ID.

C. Data Pre-processing

As the data was provided into 3 separate sheets, it had to be combined into a single file. While the other team members combined the sheets into a single table using SQL queries, I chose to take an alternative approach of performing the pre-processing task through Excel. This was done to validate and evaluate the final dataset produced by SQL and Excel for maintaining the authenticity of the data. I used Merge Tables Wizard to join the tables based on their primary and foreign key columns and the steps I followed are mentioned below:

Step 1: Select a random cell within Sheet 2 and click on 'Merge Two Tables' button on the Data Tab.

Step 2: Merge Tables Wizard will pop up, select the lookup table, and specify the column pairs to match (Model Id in Sheet 1 and Model Id in Sheet 2).

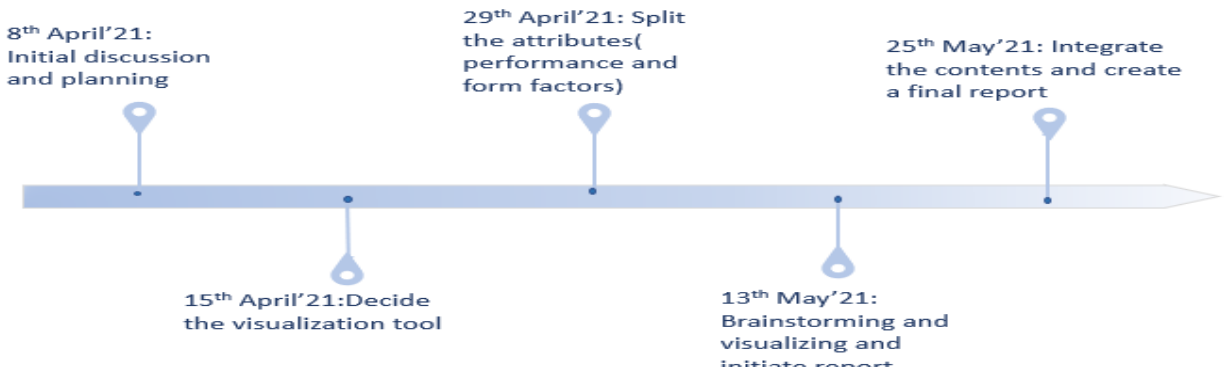


Figure 1: Timeline for the group assessment

Step 3: Choose the columns to update with the values from the lookup table and review the final result.

D. *Literature Survey*

Following my passion towards research and writing, I opted to do the Literature Survey for the group report. For the initial draft, I worked on different sections such as Abstract, Keywords, Introduction, Data types and References. I did a thorough research to determine how to distribute the attributes based on their characteristics. For instance, few journals suggested that constructs like CPU, RAM, Storage will determine the performance of a device whereas physical dimensions like Depth, Display diagonal, Mass will determine the form factors of the device.

E. *Document Management*

Team members were split into different subgroups based on the attributes they chose to work on. As far as the attributes are concerned, I took up form factors such as Display length, width, and depth to write that section in the report. Additionally, I also combined the contents contributed by other team members to create the first version of the final draft.

III. VISUALIZATION RELATED INDIVIDUAL CONTRIBUTION

This section illustrates the different visualization related tasks that I had contributed individually. It includes three main subtasks are to determine: Visualization method, Visualization Diagram, User interaction used.

A. *Determining a visualization method to be used*

1) Tool Identification: Once the single sheet of dataset was ready, we had to determine which visualization method can be used. As I have a strong hold on Tableau compared to any other tool, I chose to use this tool. For a similar reason, everyone in the group had a mutual choice.

2) Define terminologies: With so many attributes and three open ended questions, I defined the phrases such as 'new type of mobile device', 'new markets', 'different periods between 1989-2012' for better understanding in the team. I reckon, a product released by the company can be considered as a new mobile device if the performance specifications are faster and mobility is better than previous releases. For instance, iPad had larger storage capacity than other mobile devices. New markets can be identified as the market created the companies when they release a device with a feature that never existed in the market before.

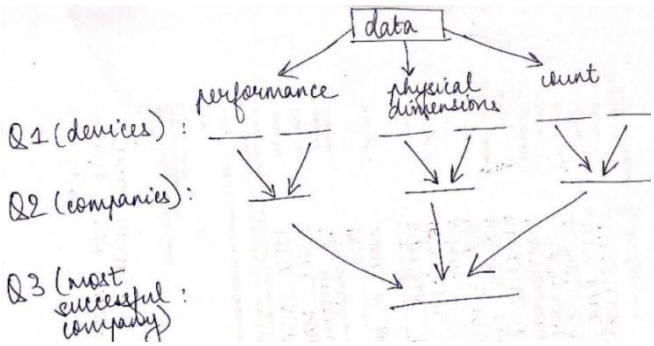


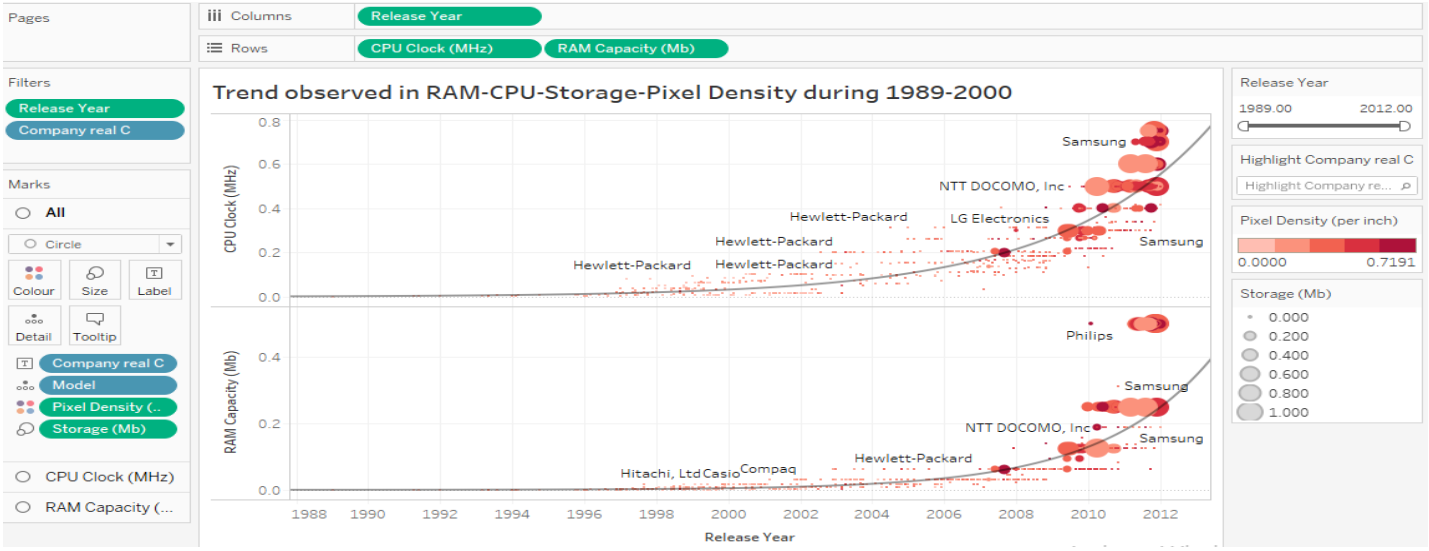
Figure 2: Tree-like structure for deriving the three answers

3) Visualization approach: I suggested this tree-like structure (Figure 2) to the team to approach the answers in a systematic manner. First the attributes are split into three categories: Performance, Form factors and Count (number of mobile devices released by the company) to answer the first question about the release of new mobile devices. On getting multiple devices with better performance and form factors, we could identify which companies tried to create new markets as an answer for question two. Once we have the list of companies leading the market, we could easily identify the most successful company leading the market (question three) by taking the count of devices released by them into account.

B. *Determining/Designing visualization (selection of axes arrangements, visual variables, etc)*

I started the visualization part by creating a bubble scatter plot (Figure 3) for four attributes: RAM-CPU-Storage-Pixel Density. This was done to observe a trend in the values of these attributes for specific companies and for a specific

Figure 3: Trend observed in RAM-CPU-Storage-Pixel Density



duration of time. A detailed explanation of the axes arrangements, visual variables and interactions used is mentioned below:

- Axes arrangement: The time period was plotted on the horizontal axis to perform temporal analysis. Whereas RAM capacity and CPU clock was plotted on the vertical axis to capture the trend in their values over a period of 1989-2000.
- Visual variables: The visual variables I used in this visualization diagram are colour, size, and text. The colour represents the pixel density of the mobile device. The higher the pixel density, the brighter the data point in the diagram. The size of the bubble represents the storage of the mobile device. The higher the storage capacity of the mobile device, the bigger the size of the bubble. The text is used to label the data points with the company's name that released the mobile device.

C. Determining/Designing user interaction, etc.

- Interactions used: Two interactions have been used to visualize the diagram. A slider is used to filter the 'Release Year' attribute. This allows us to visualize the diagram based on different time periods between 1989-2000. An interaction used in this diagram is the highlighter drop down option list for 'Company real C'. This allowed us to enhance the readability of the visualization diagram and filter out the results based on different time periods.

- Valued added by my contribution: Although, this diagram does not give direct answers to the three questions stated above, but it definitely helps in giving a head start to determine the physical attributes as our Key Performance Indicator (KPI) to find the new types of mobile devices, the companies that tried to lead the new markets, new types of mobile devices launched in the market. This diagram helped me and my team to figure out the performance trends in different companies. Also, with the help of the slider, we observed that the range of values of RAM and CPU significantly differ from 2000 and onwards.

IV. KEY CONCEPTS USED IN THE PROJECT

With the help of this unit, I learnt the basic HCI concepts, visualisation techniques and fundamental algorithms to create good and meaningful visualisations of abstract information. Many of the ke concepts covered in the lectures and tutorials helped me understand the different visualization methods.

The first step of any visualization task is to get familiar with the data. From (Week 2, lecture slide, page 7)- Data Types, I learnt to categorise each column's data type into nominal, ordinal, interval, ratio. This is a very important step and is a foundation for deciding the axes arrangements and visual variables in the subsequent steps. For instance, CPU, RAM are of ratio data type whereas Model ID is of ordinal data type.

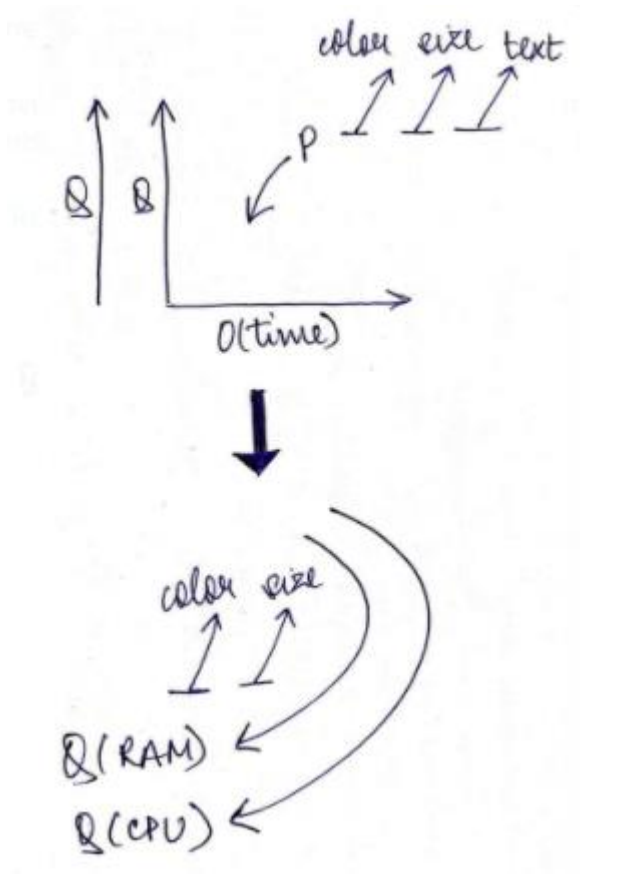


Figure 4: Symbolic Representation of Figure 3 and it's equivalent symbolic representation

Using the concept of semiology of graphics from (Week 3, tutorial, page 4), I learnt to create symbolic representations of the visualization diagrams. With the help of this symbolic representation, I created it's equivalent symbolic representation to subsequently create an equivalent visualization diagram of the original figure (Figure 4).

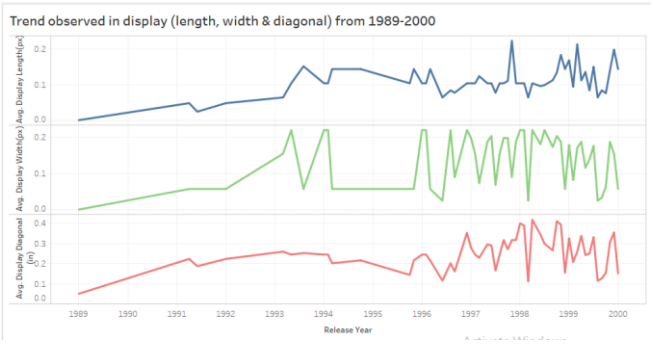
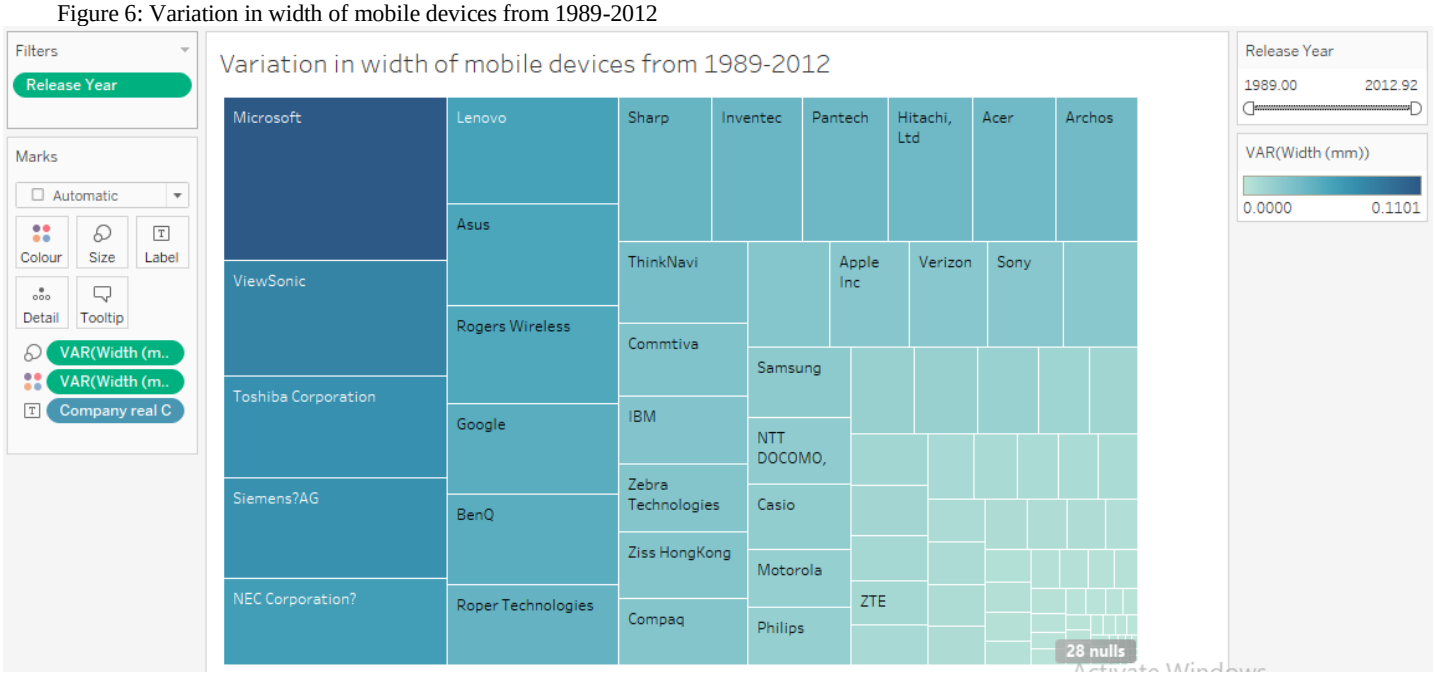


Figure 5: Trend observed in display (length, width and diagonal) during 1989-2000

Another visualization diagram that I created for this project was a trend diagram to observe the variation in display length, width and diagonal of the mobile devices. From (Week 5, lecture slide, page 7), Exploratory data analysis 1, I understood the importance of temporal analysis and this Figure 5 is the best application of this concept. A comparison is drawn between the physical dimensions of the mobile devices released over the period of 1989-2000 showing a surge in their respective values after 1992.

Figure 6: Variation in width of mobile devices from 1989-2012



One important aspect of some of the visualizations is the exploratory data analysis that can be done from statistical inferences (Week 6, lecture slide, page 3). As mentioned in the group report, the equation we used to calculate the magnitude, which is the distance for each point contains three different attributes to the origin in R3 is: $\sqrt{(CPU^2 + RAM^2 + Storage^2)}$

This equation is based on the Euclidean distance between two points [3]. The magnitude calculated can be viewed as a performance score for the model, since all three attributes can now contribute to a single value for comparison.

I created another visualization diagram to find out the companies which tried to create new markets (question 2) based on the variation in mobile width. As evident as it can be by looking at Figure 6, Microsoft has had the highest variation in terms of mobile width. This can be easily depicted by the use of colours as taught in (Week 7, lecture slide, page 7)- Visualization design and representation. The colours we wish to highlight can result in different attention span of the audience. Thus, dark blue was chosen to highlight the highest variation in width in the market.

V. Conclusion

From determining the visualization tool to final assessment of the visualization diagram, this assessment has brought out my capabilities in delivering some meaningful visualizations. Along with that, it has also enlightened me with the different non-visualization tasks, like document management and literature review which are equally important. Overall, with the help of this unit I learnt various aspects of visualization techniques, visualizing relationships, and connecting the dots to build a story. It has opened endless opportunities for me in academic research and development of new analytic methods.

References

[1] L. Collins and S. Ellis, Mobile devices. Chapman and Hall/CRC, 2015.

[2] M. Halpern, Y. Zhu and V. Reddi, "Mobile CPU's rise to power: Quantifying the impact of generational mobile CPU design trends on performance, energy, and user satisfaction", 2016 IEEE International Symposium on High Performance Computer Architecture, 2016. Accessed 10 May 2021 [Online]. Available: 10.1109/hpca.2016.7446054

[3] J. Tabak, Geometry: The Language of Space and Form, Facts on File math library, Infobase Publishing, p. 150, ISBN 978-0-8160-6876-0