

## Assignment Coversheet – GROUP ASSIGNMENT

Please fill in your details below. Use one form for each group assignment.

### Assignment Details:

Assignment Title	<b>Visual Analytics for Mobile Device Market</b>				
Assignment number	<b>2</b>				
Unit of Study Tutor	<b>Shubhi Goel</b>				
Group or Tutorial ID	<b>Tut2</b>				
Due Date	<b>May 28</b>	Submission Date	<b>May 28</b>	Word Count	<b>3301</b>

### Membership

Group Name/Number					
Family Name	Given Name (s)	Student Number (SID)	Unikey	Contribution + percentage	Signature
Du	Hongbo	500635346	hodu9417	22.5% (Q2, review and editing)	H.D.
Zhang	Wei	460024048	wzha0659	22.5% (Q3, review and editing)	W.Z.
Lin	Lin	480296188	llin4125	22.5% (Q1, review and editing)	L.L.
Syed	Ahmed Sabeeh	460498652	asye5989	22.5% (Intro, data, conclusion, review and editing)	A.S. S.
Zhao	Bingtao	490041116	bzha0935	10% (review)	B.Z.

### Declaration:

<ol style="list-style-type: none"> <li>I understand that all forms of plagiarism and unauthorised collusion are regarded as academic dishonesty by the university, resulting in penalties including failure of the unit of study and possible disciplinary action.</li> <li>I have completed the <b>Academic Honesty Education Module</b> on Canvas.</li> <li>I understand that failure to comply with the Academic Dishonesty and Plagiarism in Coursework Policy can lead to the University commencing proceedings against me for potential student misconduct under Chapter 8 of the <b>University of Sydney By-Law 1999</b> (as amended).</li> <li>This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that it is not my own by acknowledging the source of that part or those parts of the work.</li> <li>The assessment has not been submitted previously for assessment in this or any other unit, or another institution.</li> <li>I acknowledge that the assessor of this assignment may, for the purpose of assessing this assignment may: <ol style="list-style-type: none"> <li>Reproduce this assignment and provide a copy to another member of the school; and/or</li> <li>Use similarity detection software (which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking).</li> </ol> </li> <li>I have retained a duplicate copy of the assignment.</li> </ol>	
Please type in your group number here to acknowledge this declaration:	2

# Visual Analytics for Mobile Device Market

\* COMP5048 Assignment 2 - Group Tasks

1<sup>st</sup> Hongbo Du  
500635346

hodu9417@uni.sydney.edu.au

2<sup>nd</sup> Wei Zhang  
460024048

wzha0659@uni.sydney.edu.au

3<sup>rd</sup> Ahmed Sabeeh Syed  
460498652

asye5989@uni.sydney.edu.au

4<sup>th</sup> Lin Lin  
480296188

llin4125@uni.sydney.edu.au

5<sup>th</sup> Bingtao Zhao  
490041116

bzha0935@uni.sydney.edu.au

## I. INTRODUCTION

This study examines the provided data with the help of visualisation techniques and tried to uncover underlying patterns in the market. We answered the presented questions: which devices can be considered as hallmark at different time periods; which companies contributed and led the new market; and which company can be considered most successful in leading the market.

We performed various visualizations to found possible solutions, the final visualizations were chosen during meetings.

## II. INITIAL DATA EXPLORATION

The three data types in the original data set are nominal data (model, model ID, company ID, company name), ordinal data (release data, release year) and the rest attributes are quantitative data. All quantitative data were normalized.

There were some broken links in the data where the company information was not available for few devices. We managed to process and analyse the model name to derive the missing company information. The device types were also extracted from column 'model' by searching key words.

For the purpose of initial analysis and visualizing broader picture we derived two new dimensions from existing data namely "Processing Ability" which is calculated using PCA and represents CPU, RAM Storage for each device from the data set. Similarly "portability" is derived from Mass and Volume. The third dimension we considered is display which is simply represented by "display per inch" from given data. Correlation and logical relevance was considered in deriving those dimensions. These new variable are also normalized and collectively we refer them in this section as competition dimensions. We studied trend by plotting standard deviation for each year in terms of the above mentioned competition dimensions against time. The reason is that the large deviation represents variety in the market. It indicates when the trend changes, which enable us to focus point of interest.

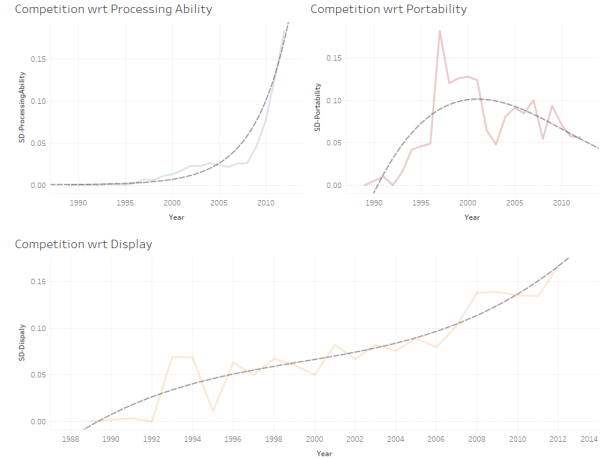


Fig.1 Competition Trend Over the Years

Figure. 1 presents the trend over the years in each competition dimensions. Reference trend lines are shown as dashed line against faded lines to data plot to bring focus to general trend. It is evident that there has been competition on processing ability, display and portability among companies. In terms of the display, there has been a gradual increase in competition and suggesting that there always been variety and competition that upped with time. It is evident that, in early years, portability was the major competition, then process ability and screen drove the competition.

## III. LEADING DEVICE

A KGI was constructed to present user experience. It uses Euclidean distance between origin and coordination of the device in  $N$ -dimensional space, where  $N$  denotes the number of variables used to indicate device specification. Some columns like width and length are redundant columns as display diagonal can present display size. They will not be used to calculate KGI. Only "Model" "Release Date" "RAM Capacity (Mb)" "Storage (Mb)" "CPU Clock (MHz)" "Display Diagonal (in)" "1- Mass (grams)" and "Pixel Density (per inch)" are used to form KGI.

The KGI has been given a name called "User experience indicator", and it's calculated by Equation (1):

$$KGI(0, q) = \sqrt{\sum_{i=1}^n (q_i - 0)^2} \quad (1)$$

where  $q$  is points in Euclidean  $n$ -space,  $q_i$  is Euclidean vectors constructed by selected columns in Table 5.

### 1) Visualisation

Figure 2 is an interactive visualization to identify which devices were trying to create new market with the new type of mobile device.

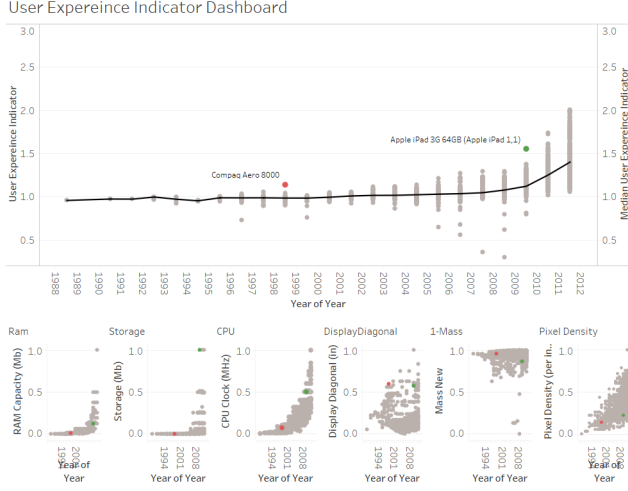


Fig.2 User Experience Dashboard

### 2) The utilization of axes arrangement

On top chart ,the x axis is year and y axis is the "User experience indicator" score. For 6 sub-charts, the x axis is year and y axis is the normalized values from original table with exception of mass. 1-Mass is used to replace Mass so 0 to 1 will represent user experience from worse to better for all the quantitative columns.

### 3) The utilization of visual variables

Each point represents one model. The 2 models which are identified as leading devices creating new market are highlighted in red and green to contrast to other devices colored by grey. The black line on top chart is the median score of all models in that year.

### 4) The utilization of user interaction

At the bottom there 6 sub-charts which represent normalized values for 6 columns we use. These 6 charts can give indication that which features are leading for top devices. The dashboard is designed to link points in main chart with the points in sub-charts so they can be highlighted at same time.

2 devices are outstanding which indicates they are trying to create new market. One device is "Compaq Aero 8000" in 1999 and another one is "Apple iPad 3G 64GB (Apple iPad 1,1)" in 2010. Fig.3 illustrates why "Compaq Aero 8000" was trying to create a new market.

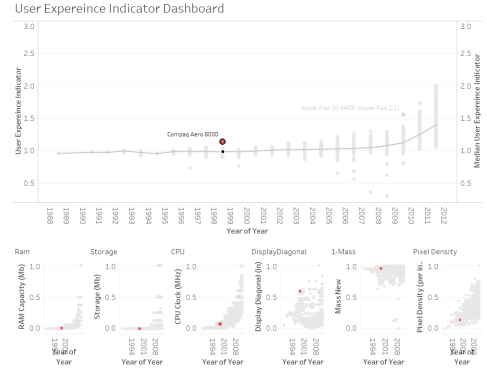


Fig.3 Compaq Aero 8000

Its score is outstanding in that year and even comparing to the devices in later years from 2000-2004, the score of this device is high. By highlighting that data point in main chart, we can see the position of this device in each feature. Here we can see its ram, storage and pixel density are just in the middle range among other devices in 1999, but its display diagonal, CPU and mass are in the top range. It seems this device was trying to create a new market by introducing large screen, high spec CPU and remain light weighted at same time.

Searching this device on internet[1], we can see the device is unique in 1999. Compaq was trying to create a new type of device mixed with notebook and mobile phone, but that idea failed because of use-ability and lack of applications.

Figure 4 illustrated why "Apple iPad 3G 64GB (Apple iPad 1,1)" was trying to create a new market in 2010.

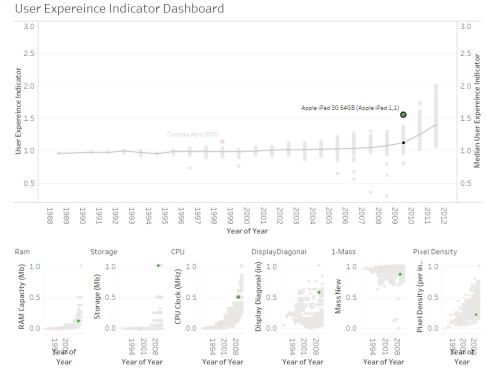


Fig.4 Apple iPad 3G 64G

The device was trying to create a new market by introducing a device with large storage, large screen, and fastest CPU. As we know in 2010, Apple iPad first generation was released, and it created a new market called pad or tablet device market [2]. The method we use here correctly picked up the leading device and is proved to be working.

## IV. COMPANIES THAT LEADING THE NEW MARKETS

### A. Top Performance Bar Plot

We produce the following bar plot to show which companies tried to lead the new markets with the new types of mobile device. Firstly, we define the new types of mobile device as tabs and pads, since those types of mobile device were

introduced later than the phone devices. According to our data cleaning and preprocessing, we successfully obtain the device type variable. We extract the top 10 companies that have the highest aggregate performance indicators of both types of devices from 2010 to 2012, and the companies are ranked from left to right in descending order. The top performance indicator, as KGI, tells the top technology at each company in terms of device performance, which is calculated as

$$I_{performance} = \max([CPU\text{Clock}(MHz)]) + \max([RAM\text{Capacity}(Mb)]) + \max([Storage(Mb)]). \quad (2)$$

Lastly, from this figure, we can directly find that Samsung, Apple, and Toshiba are the companies that tried to lead the new markets with the new types of mobile device.

### 1) Visualization

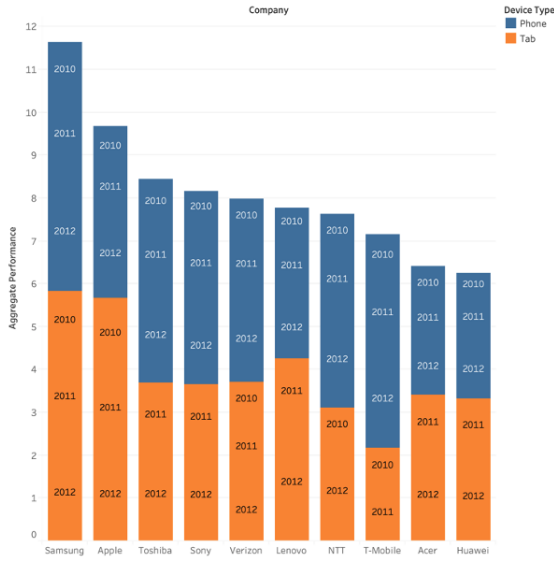


Fig 5. Top Performance Bar Plot

### 2) The utilization of axes arrangement

We arrange the company names on the horizontal axis since the company names are discrete nominal variables. It would be easier for users to see the differences between each company. We arrange the top performance indicator on the vertical axis, because the top performance indicator is ratio data. The higher bar implies the better the company is.

### 3) The utilization of visual variables

We classify the device type by orange and blue, where orange represents the tabs and blue represents the phones. The use of orange versus blue can well represent the contrast of how each company distribute the innovation on tabs and phones. Notice that there are small gaps between each bars since the companies are discrete variables. In addition, we label each year on each bar to indicate how each company innovates in every year; and the year label also points out when each company started the tab devices.

### 4) The utilization of user interaction

Comparing the top performance distribution at each company tells the users how this company innovated on the new device type. In particular, Samsung and Apple have the top tier technology in both phone and tab since those two companies have the highest aggregate performance indicator, phone and tab performance indicator as well. Companies such as Motorola and HTC do not appear on this plot since they do not have tab products good enough to be ranked in the top 10. Therefore, the users can acknowledge which companies performs well in a comprehensive aspect.

### B. Count of Device Released Heat Map

For representing the count of released device by companies, we produce a heat map. We filter the top 10 companies which released the most devices from 1993 to 2012, and we rank the companies according to the last year's release count from top to bottom in descending order. The users can directly learn that the companies such as Samsung, Sony, and LG can be considered to be the companies leading the market in terms of quantity of model invention. Since those companies released the most devices throughout years. Companies like HTC and Huawei only have a few years in records but are listed on this plot, which implies that those two companies are leading the new market as well since they are growing up fast.

### 1) Visualization

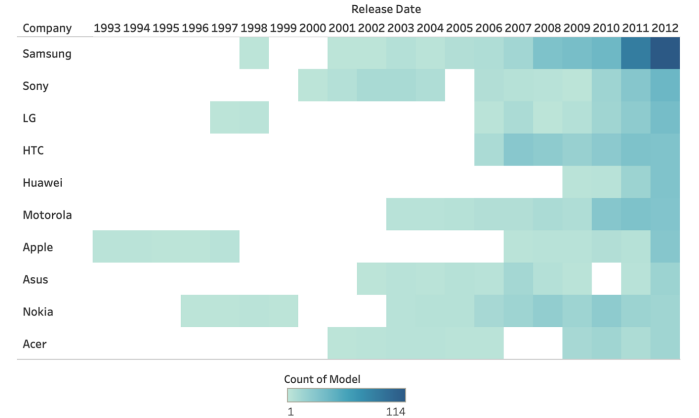


Fig 6. Count of Device Released Heat Map

### 2) The utilization of axes arrangement

We arrange the release years on the horizontal axis which ranges from 1993 to 2012; and we arrange the company names on the vertical axis. We arrange the two axes in this way since it would be easier for users to look from left to right to compare the years. Another concern is that if we arrange the years on the vertical axis, it would be hard to decide which year to be located at the bottom, because the 2012 was located at the bottom in the last figure, so it is better to keep it in the same style; however, it would look like a bar plot due to those blank blocks.

### 3) The utilization of visual variables

This heat map uses the colour hue to denote the number of models released by each company, where the darker colour hue represents the more models released by this company, and vice versa. Colour hue is the most appropriate usage since it well represents the variations in numbers. Such way of using of colour brings more users attention on the quantity of released models. We only focus on the count of released models by each company. Hence there is no need for textual labeling since the colour hue can well represent the change in quantity.

### 4) The utilization of user interaction

The users can see that Samsung released the most models in 2012, and the numbers of released models increase along with the years in the decade. So is Motorola. Those companies, such as Sony, Asus, and Acer, has one or more years that did not release the new model in a year. Huawei, HTC, and Apple are the companies that started on making phone or tab devices in this decade, and those companies invented more and more models. Therefore, these companies are considered as the leading companies in the market.

## V. THE MOST SUCCESSFUL COMPANY

### A. Capabilities of Leading Companies Pie Chart

The following pie chart is representing the proportion of leading devices that the top 10 companies own in the entire market. To find the leading devices, we set the KGI as process ability, display (DP) shape, device size and user experience. The formulas are shown below.

$$ProcessAbility = \sqrt{RAM^2 + Stprage^2 + CPU^2}$$

$$DPshape = \sqrt{DPdiagonal^2 + DPwidth^2 + DPlength^2}$$

$$DeviceSize = (1 + mass)/(1 + volume)$$

(3)

The formula for User experience indicator has been stated in section III.

We normalized each KGI based in each quarter of each year to find the ranking. We observed the top 10% devices in each KGI in each quarter of each year as the leading devices. Then we rank the frequency of devices to find out the top 10 leading companies. By compare the leading devices that owned by the top 10 companies and other competitors, the pie char below was generated.

### 1) Visualization

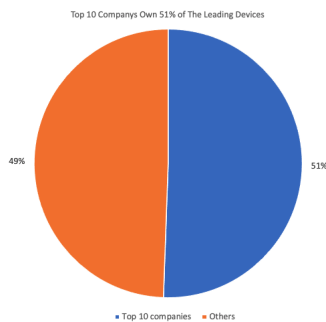


Fig 7. Top 10 companies own 51 of the leading devices

### 2) The utilization of axes arrangement

One axis in this pie chart is the percentages as quantities. Another axis could be the entity name as categories.

### 3) The utilization of visual variables

We use different colours to denote the leading companies and others. We use the arc to represent the percentage. Users are able to be aware of the leading companies' strength on non-sales related field. The percentages are labeled outside the pie chart to differentiate the similar size of pieces.

### 4) The utilization of user interaction

Users can intuitively see from the color and the arc of each sector to understand that about half of the leading devices were designed by the leading companies.

### B. Sunburst Graph to find the Most Successful Company

The following sunburst graph describes within the leading companies, how many of their products were leading in different numbers of KGI. The sunburst graph shows data hierarchically. The innermost circle represents the portion of leading devices that each company owned. The outer circle represents how many devices were leading in a number of KGIs. For example, among all leading devices of Hewlett-Packard, a small portion of them were leading in all KGIs when they were released, while Motorola only leads in one of the KGIs at a time.

### 1) Visualization

Top 10 companies that own the most number of leading devices and their frequency of occurrence In set KGIs

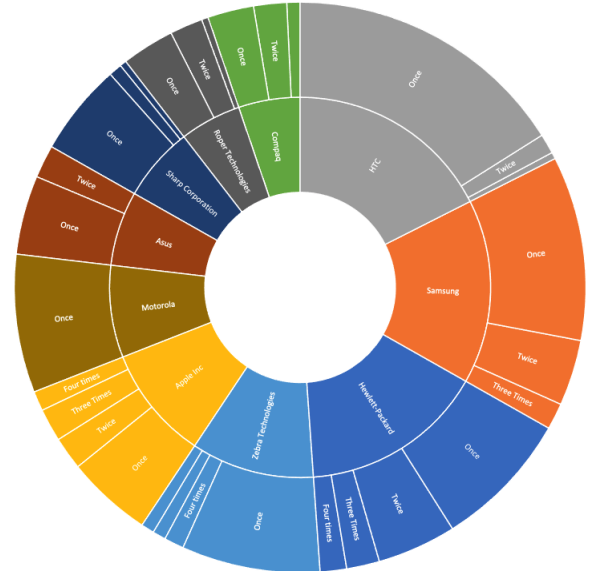


Fig 8. Top 10 percentages and frequency

### 2) The utilization of axes arrangement

The first two axes are company's name and the number of frequency as categories. The third axis is the number/percentage of the devices that companies owned.



### 3) The utilization of visual variables

Company's name and frequencies number are nominal data, the number/percentage is quantitative data. Colour hues and labels in the innermost ring were used to represent the companies, while the labels in the outer ring were used to represent the frequency number. The arc is representing the proportion of certain object.

### 4) The utilization of user interaction

This visualization is appealing to the human eyes by using colour hues. Users would first focus on the brighter part, then they will look at the grey sector by appealing to the brain with the arc size. Therefore, the top 5 leading companies were identified, they are HTC, Samsung, Hewlett-Packard, Zebra Technologies and Apple Inc.

### C. Area chart to show the top 5 leading companies' release data

The following area chart describes the release data of the top 5 leading companies that we found in the sunburst graph. The area represents the release numbers of a particular company released in each year.

#### 1) Visualization

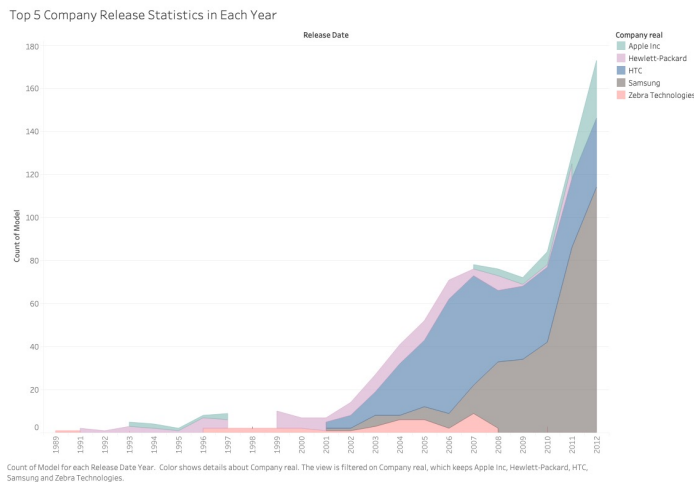


Fig 9. Number of releases by the top 5 companies

Colour grey is Samsung, Colour blue is HTC, colour green is Apple Inc, colour pink is Zebra Technologies and colour purple is Hewlett-Packard.

### 2) The utilization of axes arrangement

The horizontal axis is time, while the vertical axis is the number of mobile devices released in each year.

### 3) The utilization of visual variables

Time is ordinal data from 1989 to 2012. Number of mobile devices released in each year is quantitative data. The area represents the proportion of devices released by a particular company in each years. The colours represents the companies.

### 4) The utilization of user interaction

This visualization is appealing to the human eyes by using colour hues and areas. Users would first focus on the bigger area, then they will look at the the smaller area. Therefore, the top 2 leading companies are HTC and Samsung.

### D. The most successful company

By comparing the sunburst graph with the answer from the area chart, the other three companies and the top 2 companies have a similar number of leading devices, but the number of releases is much less than that of the top 2 companies, indicating that these three companies have higher percentages of releasing leading devices. However, it could be argued that the top 2 companies not only released leading devices but also released a large number of devices to attract mid-range and low-end users. Furthermore, Samsung has more number of devices that are leading in more than one KGIs compare to HTC. Therefore, Samsung would be considered the most successful company.

## VI. CONCLUSION

In summary, our findings are listed below.

1) Two models, Compaq Aero 8000 in 1999 and Apple Ipad 3G 64Gb, were outstanding among the competing devices of their time. They created the new markets and pushed the competitions in different dimensions.

2) Competition in terms of market lead was very close, and several companies stood out over the studied time period. Samsung and Apple were the companies that not only handled both pad and phone productions, but also were in leading positions in count of model release aspect in the new markets. Other companies like Motorola, Huawei, Sony, Asus and Acer were leading the market with the new types of devices among top companies as well.

3) Samsung is considered as the most successful company, since it introduced high-end devices to low-end devices with excellent leading rates in the overall market.

## REFERENCES

- [1] To-tech.com. 2021. Compaq Aero 8000 Windows CE mini-notebook. [online] Available at: <http://to-tech.com/windowsce/compaq/aero8000/> [Accessed 24 May 2021].
- [2] En.wikipedia.org. 2021. iPad (1st generation) - Wikipedia. [online] Available at: [https://en.wikipedia.org/wiki/IPad\\_\(1st\\_generation\)](https://en.wikipedia.org/wiki/IPad_(1st_generation)) [Accessed 24 May 2021].
- [3] "How mobile phone technology has changed over the last 40 years: Netstar," Netstar IT Support, 18-Nov-2020. [Online]. Available: <https://www.netstar.co.uk/mobile-phones-years/> [Accessed: 05-May-2021].
- [4] L. McPherson, "EVOLUTION OF MOBILE PHONE", Information Visualization, 2021. [Online]. Available: <https://studentwork.prattsi.org/infovis/visualization/evolution-of-mobile-phone/> [Accessed: 05- May- 2021].
- [5] "The History and Evolution of Cell Phones", Artinstitutes.edu, 2021. [Online]. Available: <https://www.artinstitutes.edu/about/blog/the-history-and-evolution-of-cell-phones> [Accessed: 05- May- 2021].
- [6] J. Gold, "11 tech breakthroughs that led to today's smartphones", Computerworld, 2021. [Online]. Available: <https://www.computerworld.com/article/3222855/11-tech-breakthroughs-that-led-to-todays-smartphones.html> [Accessed: 05- May- 2021].