**Information Visualization**

**Final Paper on “Line Chart visualization”**

**Department of Computer Science, Kent State University**

**CS 67302: Information Visualization**

**Bill Reed**

**Submitted by: Shiva Kumar Peddapuram**

**811235874 (**[**speddpau@kent.edu**](mailto:speddpau@kent.edu)**)**

**Abstract:**

The main intention of this paper is to summarise and give detailed information about the Line Chart Visualization. Line Chart Visualisation is the strongest visualization tool which depicts and shows the patterns and trends in each dataset over a certain time. In this paper we can explore the uses of line charts in data visualization which focuses on the strengths limitation about it. Even In this paper I am going to produce some strong evidence like strengths and weakness of this visualisation over the different datasets. This paper describes how the line charts can be used to convey useful summaries and support research conclusions.

**Introduction:**

Basically, The Line Chart is used to represent a set of information in a form of series of data points that are represented on two axes (can be of with any variable but mostly we use x and y axis) connected by a straight line. It depicts the changes in data over the time and to identify trends, patterns in each dataset. Line plots are used in plotting a wide range of the data, like stock Prices, Population growth, university students over a given period and also many more forms.

Though they look simply but gives powerful insights and are widely used in the domains of finance, economics etc.

In a line chart the horizontal axis represents mainly time or in some cases some other datasets data, and on the vertical axis represents the data that is needed to be plotted. Each data point can be represented either in dots or any other symbol (based on user interest). These set of points provides a clear visualization trend over the given time frame in the dataset.

**History of the Visualisation:**

The history of line charts can be traced back to the 18th century when they were first used to plot time series data, such as financial market data.

Different sources says that Line chart was invented by Scottish Engineer and political economist “William Playfair”. He is also the author of “The commercial and political Atlas.” Which is written in 1786. Initially he used 43-line charts variants to explore time-series data about political and economic matters.

William Playfair line chart showed the wheat price in England from 1660 to 1782 which played a crucial role in the data visualization in that time. Before the introduction of line chart data is showed in different forms like tables or simple graphs when they allowed this line chart it made easy in term so comparison for the data over different series of time and enabled users to identify different trends and patterns.

The se of line chart drastically increased in the 19th century mainly in the field of economics.

Next, the French economist Adolphe Quetelet used line charts to analyse demographics data, while the English Economist and Statistician William Stanley Jevons used line charts to study economics fluctuations.

As the 20th century started the usage of line charts became even more widespread with this technology era. Even different software’s came into to picture where it became easy to visualize charts using them.

Today line charts are commonly used in a variety of field like finance (to analyse the trends of a stock over the time), science, and engineering to visualise the data over the given period of time and identify the particular trends.

**Strengths of Line Chart Visualization:**

1. Line charts are “Easy to read and summarize” as we use simple and intuitive visual elements such as dots and lines to represent the data.
2. Line charts are the best for showing different trends in a data over the time. Making them useful foe tracking changes in data over extended periods.
3. We can use multiple datasets in a single chart making it easy to compare and analyse different trends in different datasets in a single chart.
4. The best thing about the line chart is that we can identify outliers in the given dataset by highlighting data points which are deviating drastically from the trend lines.
5. Line charts can be used to represent the carious data types of data like continuous categorical data making flexible tool for data visualization.
6. Line charts are effective for presentations and reports. As they are effective in communicating trends.
7. Line charts can help decision-makers identify trends and make informed decisions based on data-driven insights.
8. Line charts can be used for larger data sets which make them useful for analysing larger trends over extended periods.

**Weakness of the Visualization or Limitations:**

1. Line charts are limited to one dimension that means they can display the data along one dimensions. If we want to compare multiple variables, we need to use multiple line charts.
2. Line charts are limited to continuous data.
3. If we have multiple lines on a chart, it gets difficult to compare them accurately. Mainly when they have different scales or units of measurements.
4. If the datasets have a large set of data, then it may hide the data. This can be rectified by increasing or managing the size of the chart.

**Visualization:**

Dataset from Kaggle: <https://www.kaggle.com/code/mrwolfgang/google-playstore-eda>

Example of the Line chart Visualization using Google Play store dataset:

I have downloaded the dataset from the Kaggle the link for the dataset has been provided above and then I have opened the file in tableau desktop and then cleared all the null values and then loaded the data into MS-Excel.

Now I have taken 6 objectives and then performed different operations on it I got line chart visualizations.

**Visualization-01:**

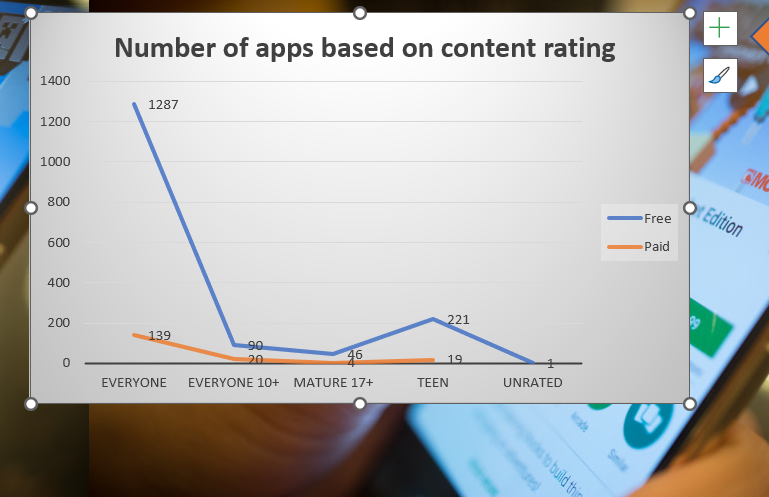
In the below visualization we can see the top 10 apps with high review count. The number on the point shows the number of reviews for app that means for angry birds epic RPG it has a review count of 2634605 in this way it follows for Top 10 Apps

Chart, line chart

Description automatically generated

**Visualization-02: (line chart with multiple parameters)**

In the following visualization I have made a line chart visualization and I have compared number of apps based on content rating like is a particular app suitable for everyone or else is it restricted to any age group and here we have two points that is free version and paid version of the app.



Visualization-03 (number of apps that are most recently updated in finance category)

A picture containing timeline

Description automatically generated

Chart

Description automatically generated with medium confidence

In the above visualization we can see the number of apps that are most recently updated that is monthly wise apps that are updated.

I have also added numbers on the top so that it becomes easy for us to determine number of updates.

Reference:

History:

1. <https://infogram.com/blog/the-line-chart-how-and-when-to-use-it/#:~:text=Many%20sources%20give%20credit%20for,about%20political%20and%20economic%20matters>.
2. <https://www.kaggle.com/code/mrwolfgang/google-playstore-eda> (for the dataset)