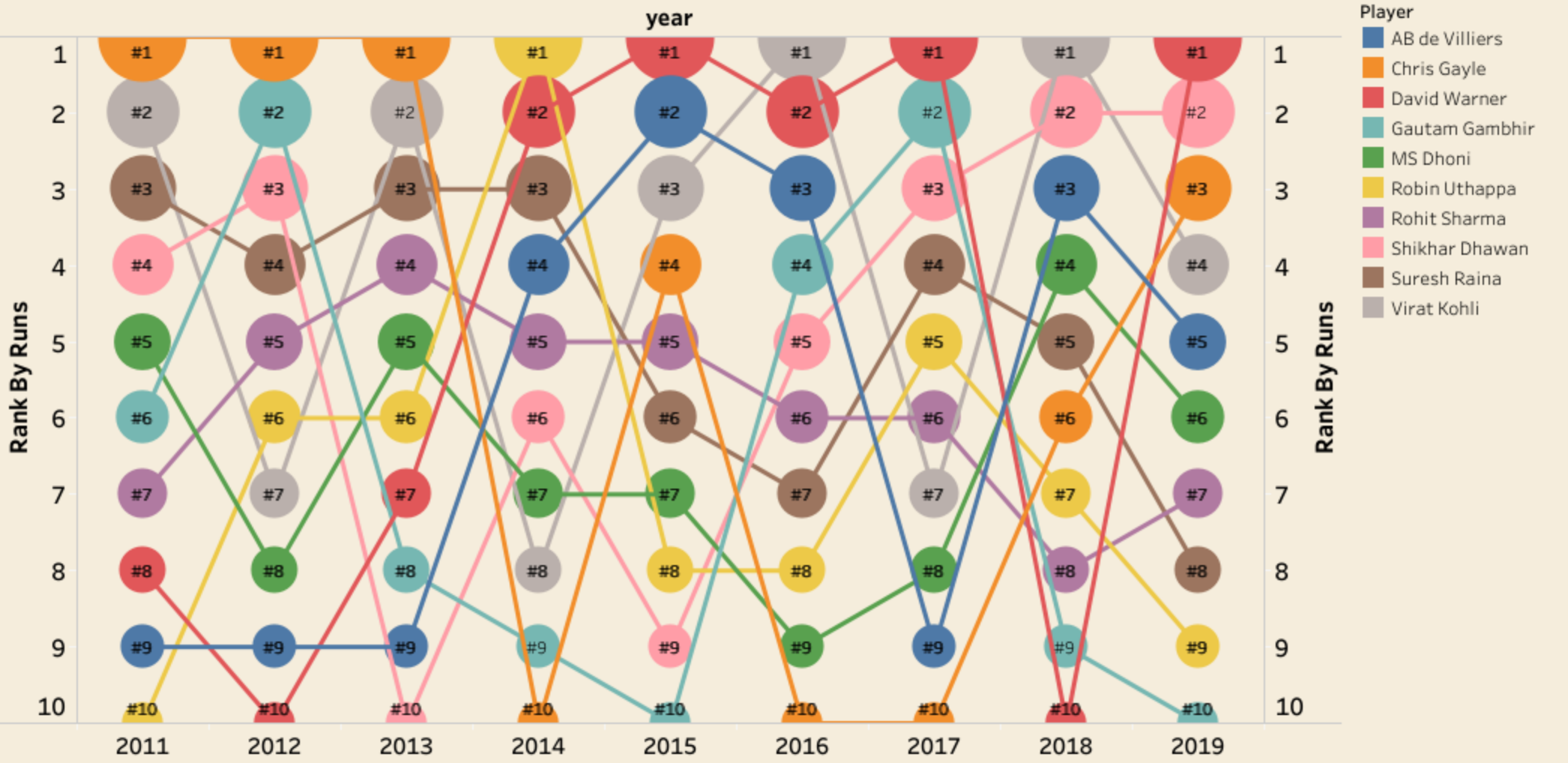
Temporal Charts

Date: 22/04/2020 **V.Saikiran(194161016)**

- Charts from this family are primarily used to show trends, events/actions or activities at a set of time-intervals. Let us discuss a few examples that belong to the temporal family. The charts include:
    1. Bump Chart
    2. Slope Chart
    3. Area Chart
    4. Line Chart
    5. Stacked Area Chart
    6. Stream Chart
    7. Gantt Chart
    8. Connected ScatterPlot
- For the sake of demonstration we explore different data collected from different sources and we use tools like **Tableau(a visualization tool),R ggplot, Altair library(Python)** which are publicly available.I will mention the data source, tools used in the respective charts.

Bump chart demonstrating IPL season wise standings of top players



Bump Chart:

- Bump chart is an alternative to Line chart, when we want to show how quantitative values, in terms of ranking measurement, have changed over time for different categorical items. In this case, we keep rank on the Y-axis and X-axis represents a continuous time scale.

Data and Data source:

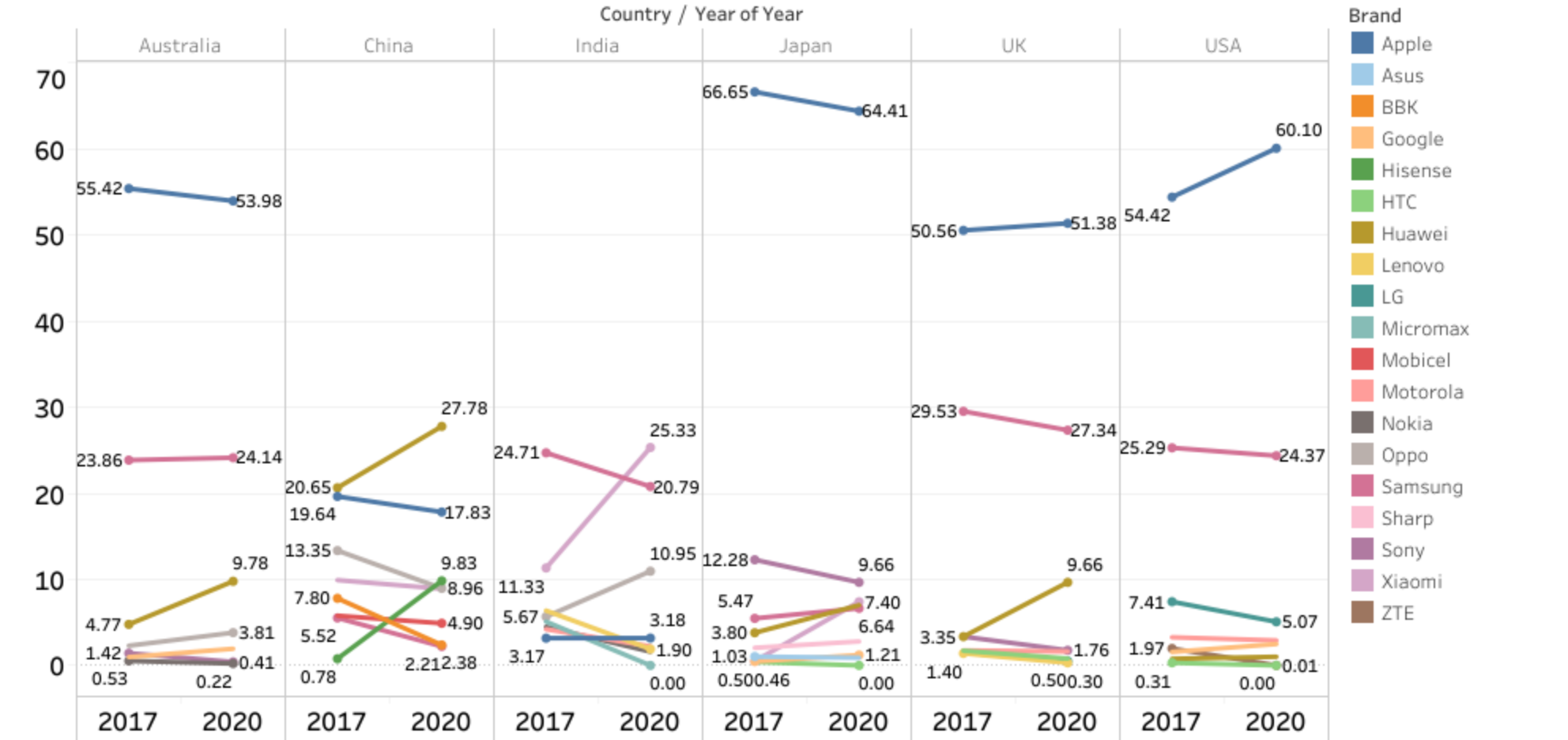
- To showcase this chart I have scraped **ipl data** consisting of **runs scored by top 10 batsmen season wise(2011-2019).**
- Source : http://iplt20.com/stats

Tool Used: Tableau

Analysis:

- Interesting thing we can observe in this chart is **David Warner** is consistently performing and making to **#1 or #2** from 2015.We can also see that in the year of 2018 he got banned from ipl due **ball tampering** issue. We can say that he is one of the consistent players.
- If we look at **Chris Gayle,** he performed well in the earlier seasons and made up to #1 till 2013. After that we can see that there is a bit of degradation in his performance.
- If we look at our captain **Virat Kohli,** we can observe a pattern of oscillations season wise. He is not keeping a consistent performance across the seasons.
- We can also see that from 2015 **AB De villiers** is also showing up a sort of oscillatory performance same as Kohli. As they are from the same team and both of them try hard to win the game with their partnership, we can expect this kind of pattern.
- **Rohit sharma** is also consistently performing well but could not score more runs which made him to get satisfy with **#4-#8** spots.
- In this way Bump charts are very much useful when we want to **rank across categories/ objects and compare the ranks over the time**.
- By using this we can observe for the general patterns to look for observations such as consistent trends **(largely parallel lines)** or completely non-relational patterns **(lines moving in all directions)**.

Slope Graph of Market Shares of different brands in different countries



Slope Graph:

- Slope Graph is used to show **trends of change of quantitative values** during the **two time points** for different category items.
- The graph consists of two parallel quantitative axes with a common value range.
- A line is plotted connecting the two axes together with the vertical position on axes representing the respective quantitative values.
- These connecting lines form slopes that indicate the upward, downward, or stable trend between the two temporal axes.

Data & Data Source:

To demonstrate this graph I have collected data consisting of the top mobile brands, its market shares in different countries. I have collected data for the year 2017 and March 2020 for **Australia, China, Japan, India, USA, UK.**

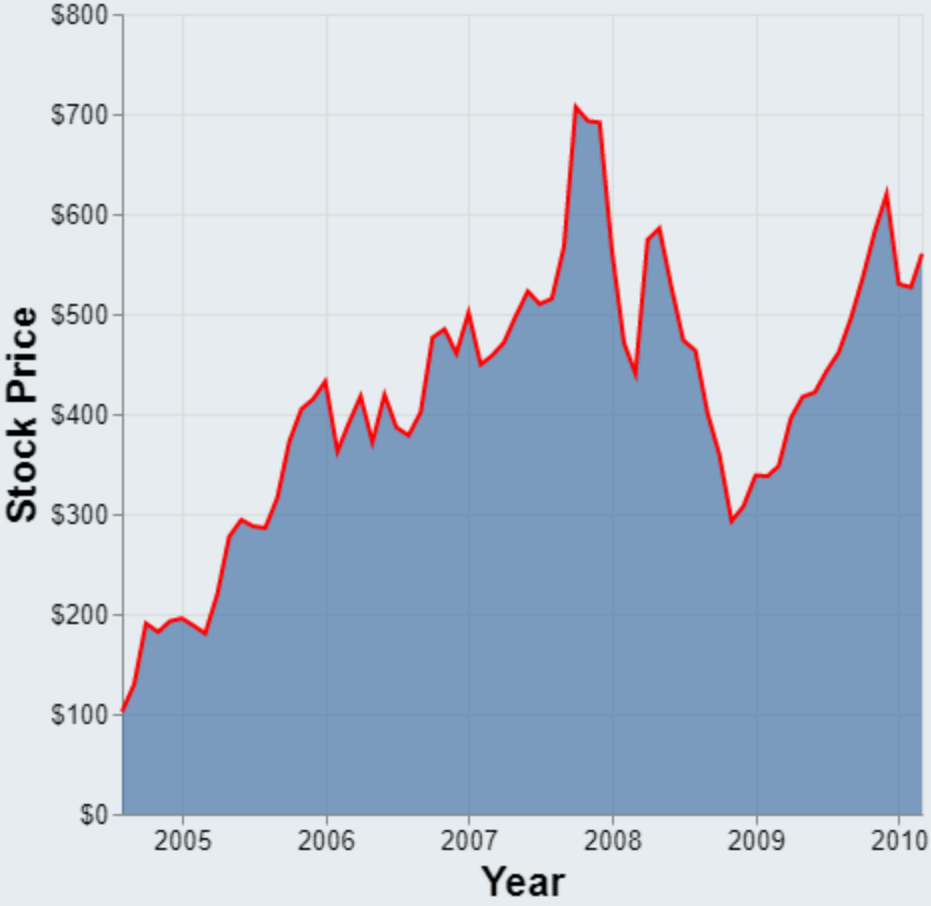
Source: **MobileBrandsMarketShare**

Analysis:

- We can clearly see that **Apple** is the leading brand with the highest market share in 4 countries of our consideration. It is not only leading but also has a **huge margin over other brands** and holding more than **50%** of market share.
- If we observe why it is lagging behind in **India**, **China** we can say that the two countries are **hugely populated** and majority of the population use **low cost mobile brands** like **Xiaomi, Lenovo, Samsung(some models out of huge no. of models) , vivo, oppo, micromax etc**.
- Apple has seen a bit increase in its share in the UK, USA and a drop in the rest of the four countries.
- After **Apple,** we can observe that **Samsung** is the second leader in most of the countries and we can say this because of its huge no. of models which are made available for all classes(wealth wise) of people.
- We can see that its share in the market is getting dropped in all the countries.
- Especially in **India**, **Xiaomi** has seen a good increase in its market share crossing **Samsung**.

Area chart demonstrating the stock price of Google

data showing for the period 2005-2010



Area Chart:

Area chart is a sort of line chart, with the area between the x-axis (representing time) and the line (quantitative measure).

Data & Data source:

For the demonstration of area chart I have used the stock price of **Google** from the year 2005 - 2010.

Source: data is available in **vega datasets** as a python package.

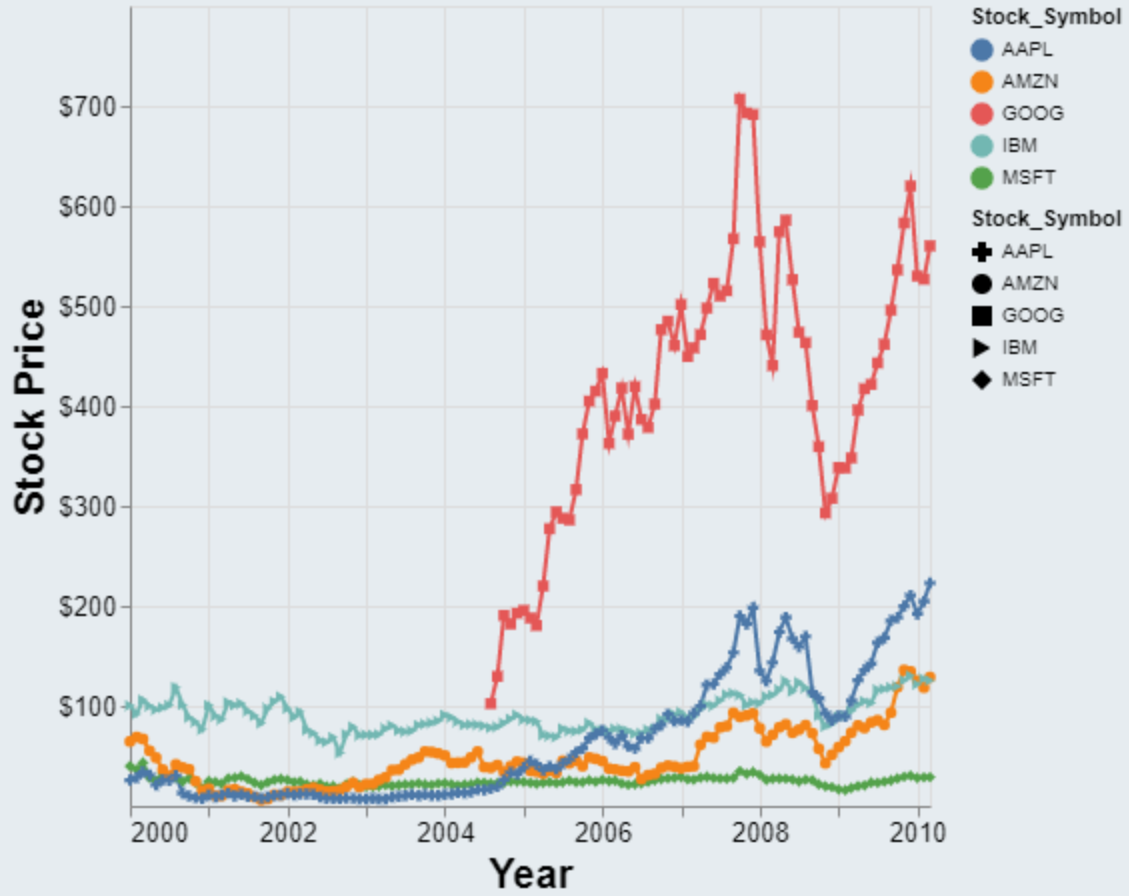
Tool Used: **Altair package** available in python.

Analysis:

- We can observe that the minimum value of Google stock price is around $100 and maximum is more than **$700**.
- Right from the year 2005 the stock price is consistently increasing and reached a peak price more than **$700** during the period **2007-2008**.
- We can see during the year **2008-2009 Google's stock price** has seen a huge downfall which was due to the **financial crisis** that happened worldwide in the year **2008**.
- The financial crisis was primarily caused by **deregulations** in the **financial industry**.
- Not only **Google** but also many companies had seen a huge downfall and experienced huge losses during the year 2008-2009.
- After that period **Google** has become a leader as an internet company with many of its widely provided services like **cloud, video sharing, maps, browser etc.**

Line chart demonstrating Stock prices of top five companies

Data showing for the period 2000-2010



Line Chart :

- A line chart displays **how quantitative values have changed over time for different categorical items.**
- A line chart is drawn around a continuous **temporal x-axis** and **quantitative y-axis** with values plotted at relevant coordinates.
- Connecting lines join up adjacent and related categorical items to form slopes which are then extended along the full timescale.

Data and DataSource: For the demonstration of line charts I have used the stock prices of top companies like **Apple, Amazon, IBM, Microsoft, Google** for the period 2000 - 2010.

Source: Data is available in vega datasets as a python package.

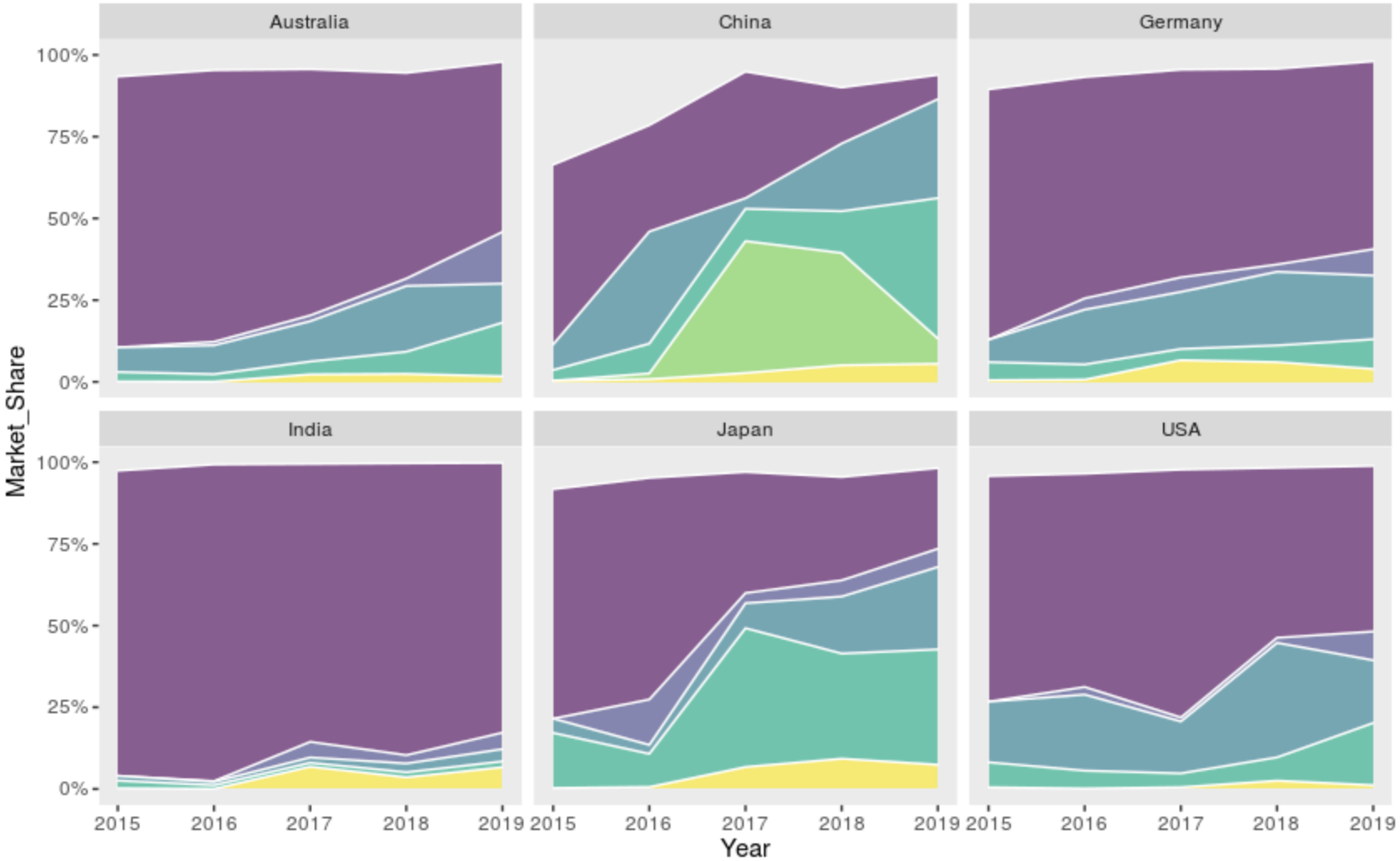
Tool Used : **Altair package** available in python.

Analysis:

- From the graph we can observe that the minimum stock price of all the companies is somewhere around **$40**.
- Looking at the chart our focus will definitely shift to the **red line** which is **performing extremely well** among all the companies.
- We can see that the red line is none other than **Google.** We can say that Google has gained huge popularity because of its wide variety of services surpassing huge tech giants like **Microsoft, IBM.**
- It is not only leading but having at least a margin 2-3 times that of other companies.
- We can see that during the period **2008-2009** all the companies experienced the **downfall** in their respective stock prices due to the financial crisis.
- **Apple** has started the **lowest** and consistently improved over the time.
- If we observe the stock price of **Microsoft,** during the period it consistently remained under $100, being the company with the **least price** in the list.
- We can also see that none of the company's stock prices except **Google has** crossed the $300 mark during the period 2000-2010.

Stacked area chart demonstrating market shares of top social media companies

Plotted country wise, 2015-2019



Data source: http://gs.statcounter.com

Stacked Area Chart:

- Stacked area chart is an area chart displaying **how quantitative values have changed over time for multiple categorical items.**
- The idea is very similar to **stacked bar chart** and allows visualization of evolution of both **total** and **contribution** of **component sub-categories**.
- However, if there are too many subcategories, then one can think of using Area Chart with faceting.

Data and Data source:

For this chart I have scraped the data of top social media companies like **Facebook, Pinterest, Twitter, Youtube, Instagram, Youku(Chinese company)** and their market share in different countries during the period **2015-2019.**

Source: collected data from gsStatcounter.

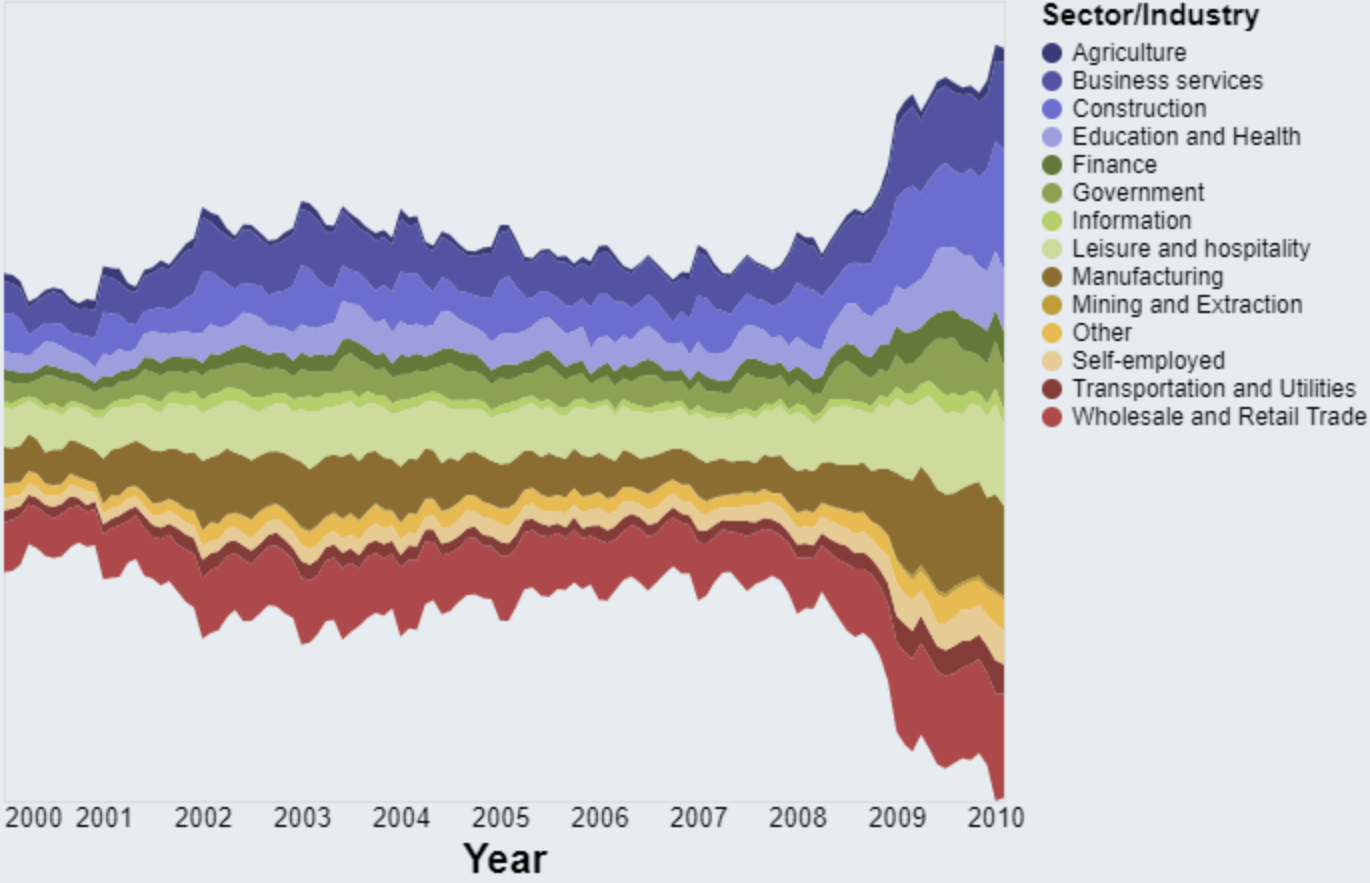
Tools used: GGplot(R, R studio)

Analysis:

- Firstly we can see that in 4 countries(Australia,India, Germany, USA) **Facebook** is the leading social media company with its highest share in **India**(~90%).
- If we look at the data of China, **Youku** a chinese video sharing company has gained a huge popularity in China during the period 2017- 2018 and gradually declined during the period **2018-2019.**
- Another thing we can observe in **China** is Facebook is losing its market share and majority of the users are showing interest in using **Instagram,Pinterest, Twitter.**
- Coming to **Japan, Twitter** has gained a huge jump in its users during the period 2016-2017 and consistently maintaining the user base.It is the leading social media company in Japan as per the statistics collected from the website.
- **Facebook** is also losing its market share in **Japan** since 2018.
- **Youtube** is falling behind all the other companies all over the period.
- In the early stages like 2015- 2016 in **India**, **Facebook** was holding more than 95% of the market share due to its popularity amongst the adults.
- In **Australia,** even though Facebook was holding a majority of the market share, it gradually is decreasing over time due to many other applications like Instagram, Twitter etc.

Stream graph demonstrating unemployment across industries

Data showing for the period 2000-2010



Stream chart:

- A stream graph shows how quantitative values for different categories have changed over time.
- They are generally used when you have many constituent categories at any given point in time and these categories may start and stop at different points in time
- The stacking arrangement of the different categorical streams goes above and below the central axis line to optimise the layout but not with any implication of polarity.

Data and data source:

For this I have used the **unemployment data across different industries** during the period 2000-2010.

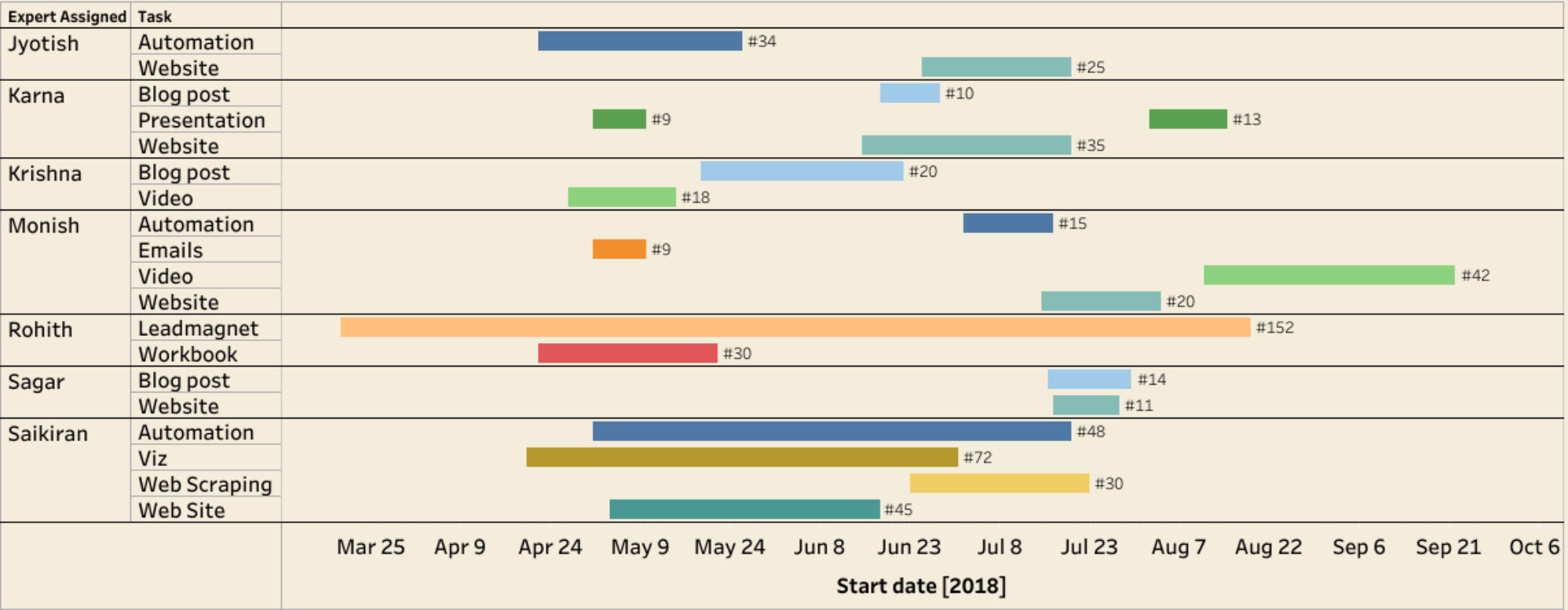
Source: Data is available in vegadatasets as python package.

Tool used: Altair package available in python.

Analysis:

- In most stream graphs we do not see the **quantitative y-axis** scale because the level of reading is more about getting a gist for the **main patterns in a relative sense** rather than an **absolute one.**
- Firstly, we can observe that there is huge unemployment across all the industries during the year 2008-2010, because of the financial crisis occured during that period as discussed in the previous charts.
- We can observe that among the industries **Agriculture, Business services, Leisure & hospitality, Transportation & utilities, and Wholesale & retail** industries faced huge unemployment.
- **Government, Information, Finance** had comparatively less unemployment during the period 2008- 2010.
- Across the industries there is a rise in **unemployment** from the year **2001-2004** as well.
- Among the industries **Finance and Information industries** have seen very less unemployment all along the period.

Gantt Chart demonstrating a project and its activites over a period of six months



Gantt Chart:

- Gantt chart is used to visualize **sub-tasks/sub-events of an event or a project over time.**
- Typically, **x-axis is used to represent the continuous time intervals and sub-tasks/sub-events on the y-axis.**
- Corresponding to each item on the y-axis, size (length) of a **horizontal bar indicates the duration of the activity**.
- Horizontal bars corresponding to the different activities may have different start and end time points. It can be further extended to visualize dependencies among several sub-tasks.

Data and Data Source: For this graph I have created a **toy data** which has different **sub tasks** related to a project over a period of **six months**. It also has data of people who are working on the subtasks respectively.

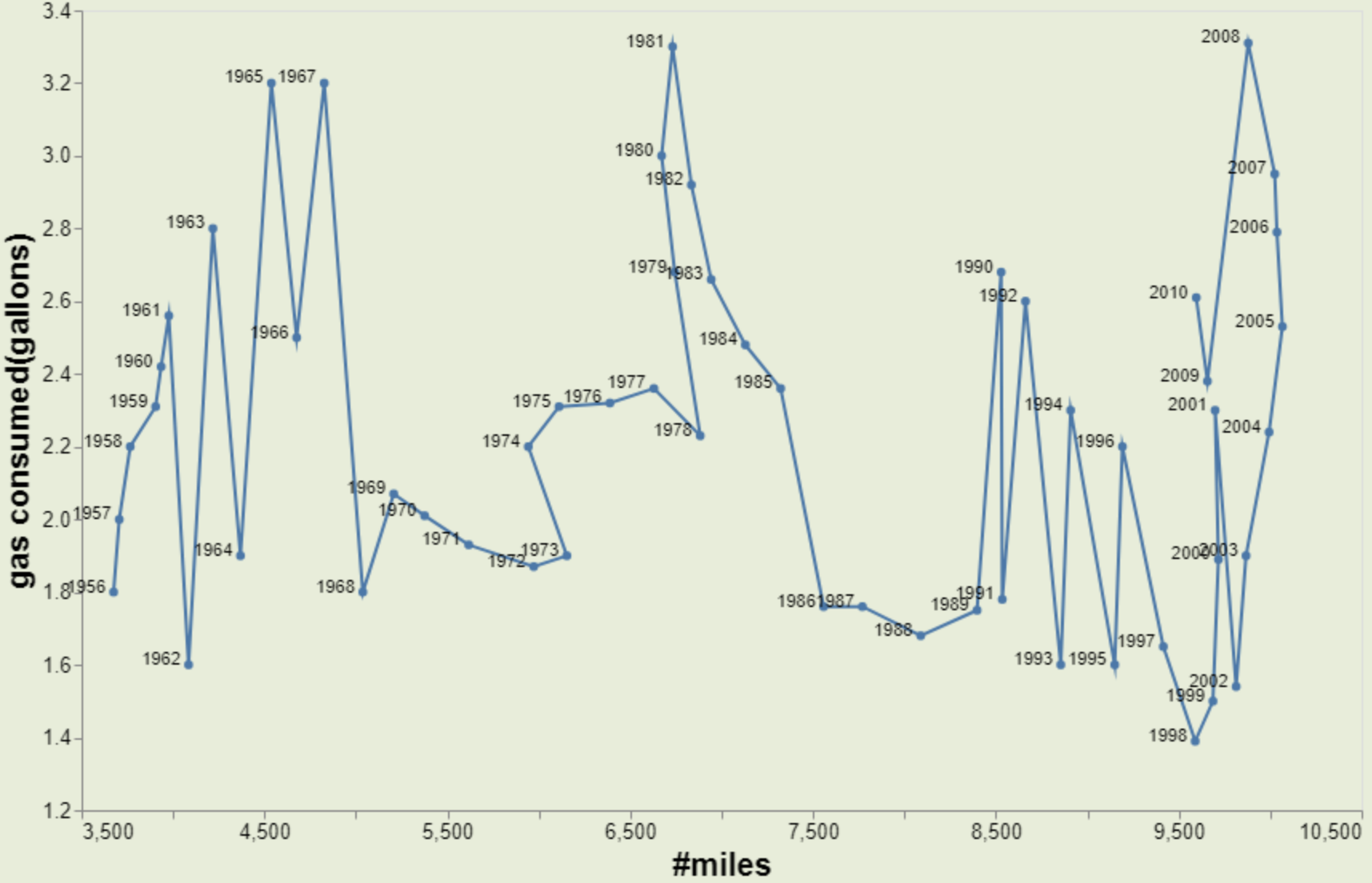
Tool used: Tableau

Analysis:

- We can observe that there are a set of 7 people working across 11 different subtasks of the project.
- Out of those Rohith has to do a lot of work on lead magnet for about 152 days.
- A **lead magnet** is a marketing term for a free item or service that is given away for the purpose of gathering contact details; for example, **lead magnets** can be trial subscriptions, samples, white papers, e-newsletters, and free consultations.
- Amongst all the subtasks, **Monish** has to do very less task for 9 days on **emails.**
- We can easily see that **saikiran**, **Monish** has got more no of tasks, but Saikiran's tasks will take up more time than Monish's tasks.
- Sagar has to work on a **Blog post, website** for about **25 days** with the least amount of tasks assigned to a person in the team.
- Amongst the task **Website** task is given to 4 people(highest) then followed by **Automation.**

Connected Scatterplot demonstrating Gas consumed against no. of miles covered

Data showing for the period 1956-2010



Connected Scatter plot:

- A connected scatter plot **displays the relationship between two quantitative measures over time.**
- The display is formed by plotting marks like a dot or circle for each point in time at the respective coordinates along two quantitative x- and y-axes.
- The collection of individual points is then connected (think of a dot-to-dot drawing puzzle) using lines joining each consecutive point in order of time to form a sequence of change.

Data and Data source: For this I have used the data consisting of no of miles covered(consecutive) and no of gallons of gas(*1000) consumed by a family.

Data source: vega datasets.

Tool: Altair python package.

Analysis:

- From the chart it is clear that the gas consumed is oscillating over the time.
- The highest consumption is during the years 1981, 2008 which was around (3300 gallons) and the lowest consumption was during 1998 which was around 1300 gallons of gas.
- We can see that there is a gradual increase in consumption during the period **1973 - 1981.**
- After **1981,** there was a sharp fall in consumption until 1991 and then the pattern of oscillation continued.
- This chart usually helps to parse our thinking by considering what **higher/lower** values mean for each **quantitative axis individually** and **then combining the joint meaning thereafter.**