

UNITEDWORLD SCHOOL OF COMPUTATIONAL INTELLIGENCE (USCI)

Summative Assessment (SA)

Submitted BY

Niraj Sharma

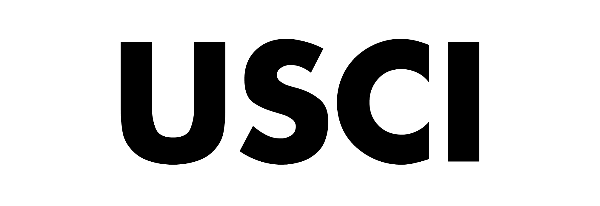
(Enrl. No.: 20220701012)

Course Code and Title: 21BSCS23C02

R programming

B.Sc. (Hons.) Computer Science / Data Science / AIML

III Semester – July – Nov 2023



Nov/Dec 2023

Contents

[Statistical Analysis and data visualization of Iris Dataset in R. 3](#_Toc152690554)

[Dataset used: Billionaires Statistics Dataset (2023) 3](#_Toc152690555)

[Objective: 4](#_Toc152690556)

[Performing the functions:- 5](#_Toc152690557)

[Statistical analysis 5](#_Toc152690558)

[ Dim 5](#_Toc152690559)

[ Head 5](#_Toc152690560)

[ Tail 5](#_Toc152690561)

[ Summary 5](#_Toc152690562)

[ Nrow 5](#_Toc152690563)

[ Ncol 5](#_Toc152690564)

[ Names 5](#_Toc152690565)

[Top 20 – by rank 8](#_Toc152690566)

[Data visualization 9](#_Toc152690567)

[Pie chart - 1 9](#_Toc152690568)

[Pie chart - 2 10](#_Toc152690569)

[Pie chart -3 10](#_Toc152690570)

[Histogram 12](#_Toc152690571)

[Bar graph 13](#_Toc152690572)

[Top 20 by tax rate 15](#_Toc152690573)

[Map graph 16](#_Toc152690574)

# Statistical Analysis and data visualization of Iris Dataset in R.

## Dataset used: Billionaires Statistics Dataset (2023)

**Introduction:**

This dataset contains statistics on the world's billionaires, including information about their businesses, industries, and personal details. It provides insights into the wealth distribution, business sectors, and demographics of billionaires worldwide.

**Structure of the dataset:**

There are a total of 2460 observations in this dataset with 35 columns

It includes :

* rank: The ranking of the billionaire in terms of wealth.
* finalWorth: The final net worth of the billionaire in U.S. dollars.
* category: The category or industry in which the billionaire's business operates.
* personName: The full name of the billionaire.
* age: The age of the billionaire.
* country: The country in which the billionaire resides.
* city: The city in which the billionaire resides.
* source: The source of the billionaire's wealth.
* industries: The industries associated with the billionaire's business interests.
* countryOfCitizenship: The country of citizenship of the billionaire.
* organization: The name of the organization or company associated with the billionaire.
* selfMade: Indicates whether the billionaire is self-made (True/False).
* status: "D" represents self-made billionaires (Founders/Entrepreneurs) and "U" indicates inherited or unearned wealth.
* gender: The gender of the billionaire.
* birthDate: The birthdate of the billionaire.
* lastName: The last name of the billionaire.
* firstName: The first name of the billionaire.
* title: The title or honorific of the billionaire.
* date: The date of data collection.
* state: The state in which the billionaire resides.
* residenceStateRegion: The region or state of residence of the billionaire.
* birthYear: The birth year of the billionaire.
* birthMonth: The birth month of the billionaire.
* birthDay: The birth day of the billionaire.
* cpi\_country: Consumer Price Index (CPI) for the billionaire's country.
* cpi\_change\_country: CPI change for the billionaire's country.
* gdp\_country: Gross Domestic Product (GDP) for the billionaire's country.
* gross\_tertiary\_education\_enrollment: Enrollment in tertiary education in the billionaire's country.
* gross\_primary\_education\_enrollment\_country: Enrollment in primary education in the billionaire's country.
* life\_expectancy\_country: Life expectancy in the billionaire's country.
* tax\_revenue\_country\_country: Tax revenue in the billionaire's country.
* total\_tax\_rate\_country: Total tax rate in the billionaire's country.
* population\_country: Population of the billionaire's country.
* latitude\_country: Latitude coordinate of the billionaire's country.
* longitude\_country: Longitude coordinate of the billionaire's country.

# Objective:

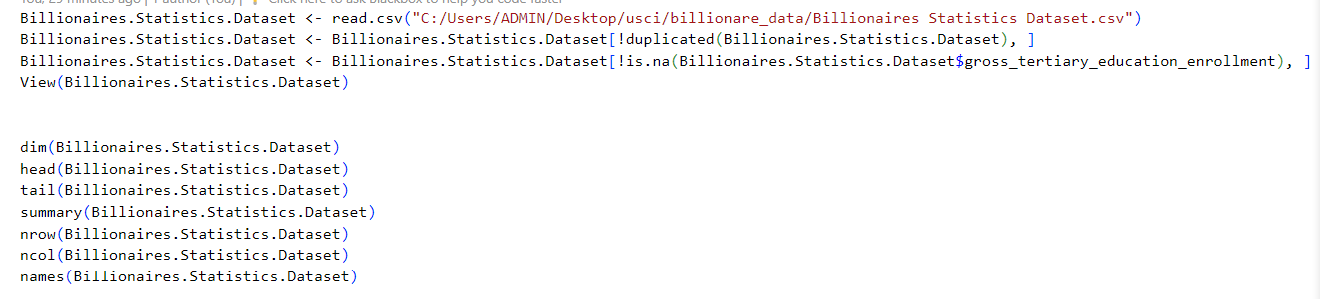
* Wealth distribution analysis: Explore the distribution of billionaires' wealth across different industries, countries, and regions.
* Demographic analysis: Investigate the age, gender, and birthplace demographics of billionaires.
* Self-made vs. inherited wealth: Analyze the proportion of self-made billionaires and those who inherited their wealth.
* Economic indicators: Study correlations between billionaire wealth and economic indicators such as GDP, CPI, and tax rates.
* Geospatial analysis: Visualize the geographical distribution of billionaires and their wealth on a map.
* Trends over time: Track changes in billionaire demographics and wealth over the years.

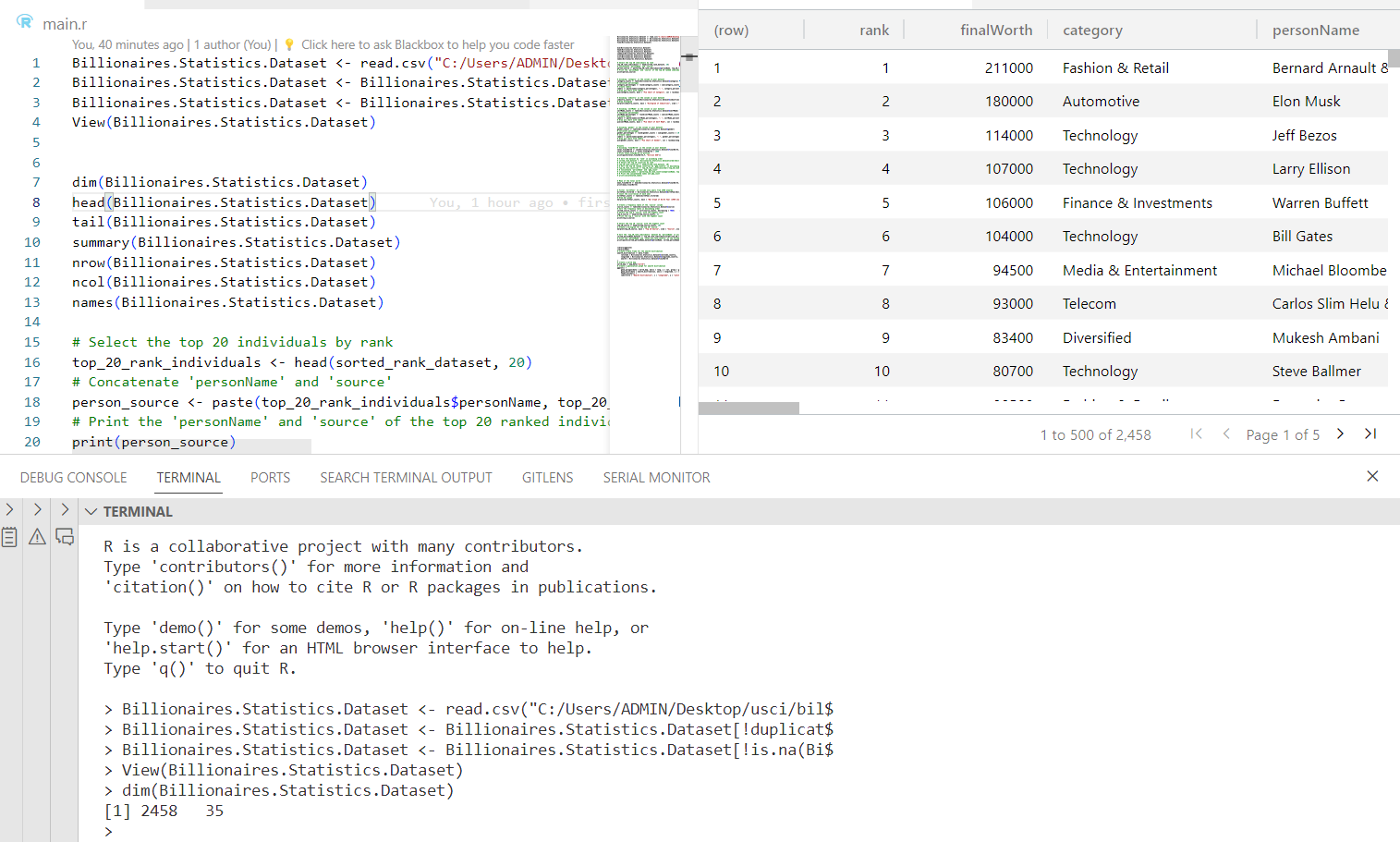
# Performing the functions:-

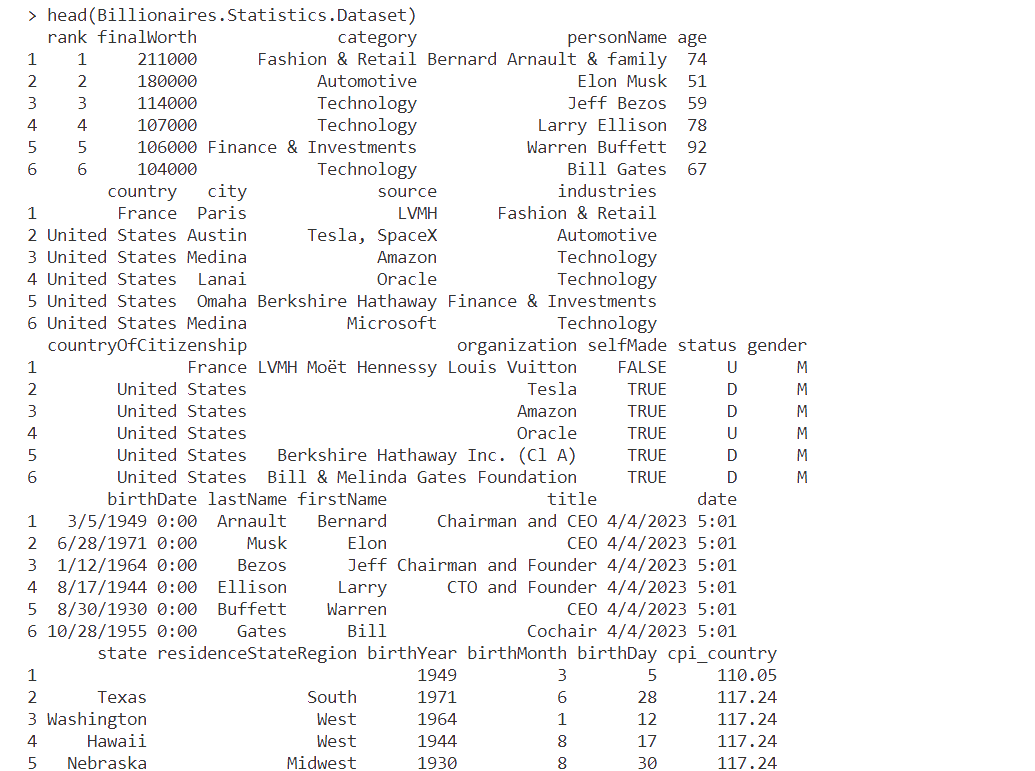
## Statistical analysis

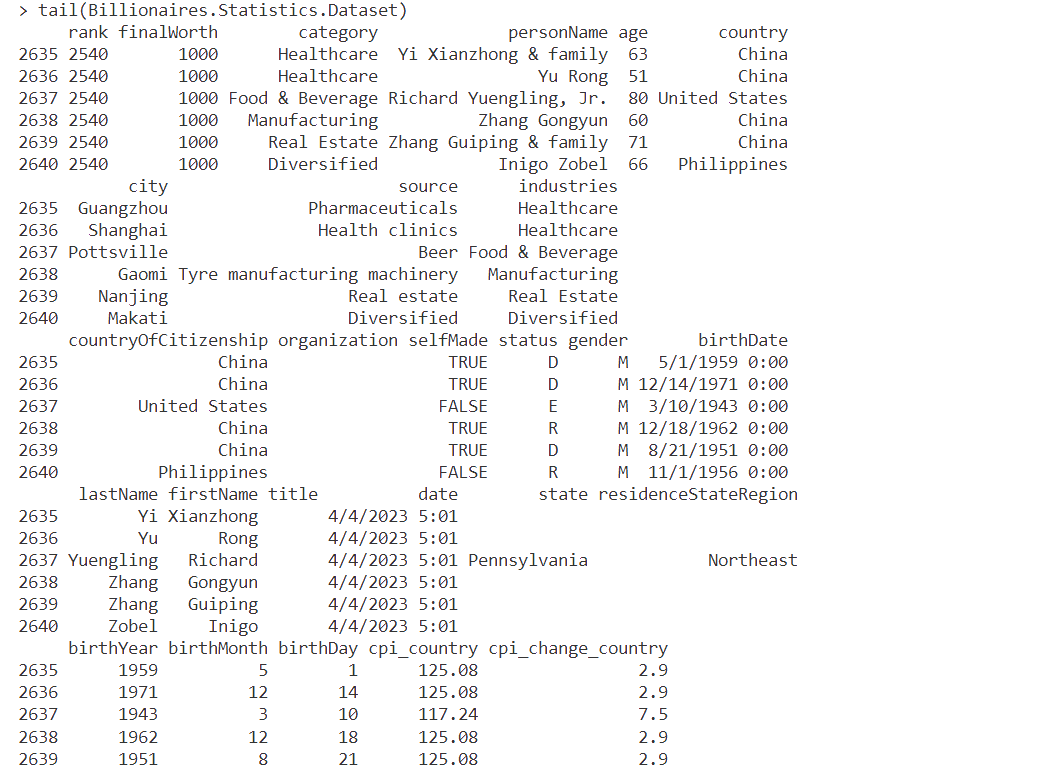
Importing the dataset and cleaning it off the duplicates and removing the entries with the NULL data in gross\_tertiary\_education\_enrollment to make it easy for further statistical operations. The View function is to view the whole dataset.

* Dim - Function is used to retrieve or set the dimensions of an object.
* Head – To display the first few rows of a data frame or a matrix.
* Tail - Function in R is used to display the last few rows of a data frame or a matrix.
* Summary - Function in R is a versatile tool for obtaining a summary of the central tendency, dispersion, and distribution of a dataset.
* Nrow – To get the number of rows.
* Ncol – To get the number of columns.
* Names - Function is used to retrieve or set the names of the columns in a data frame or the names of the elements in a list.



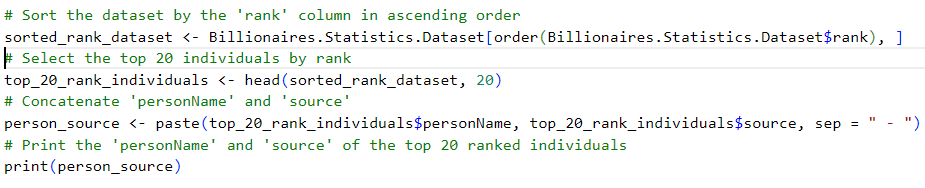


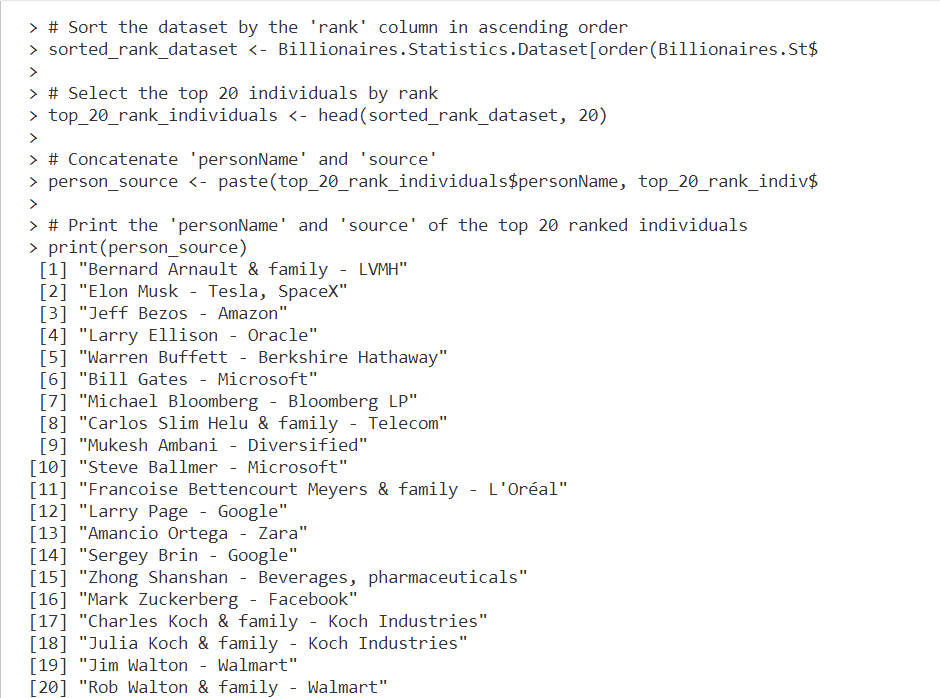






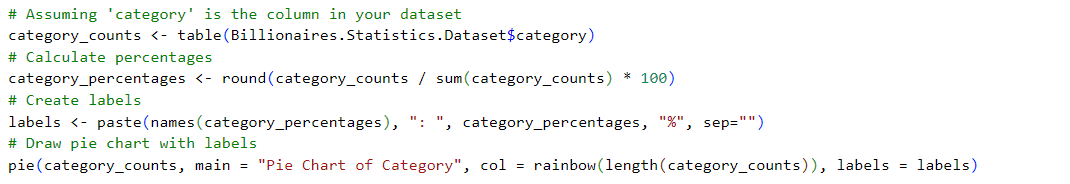
## Top 20 – by rank

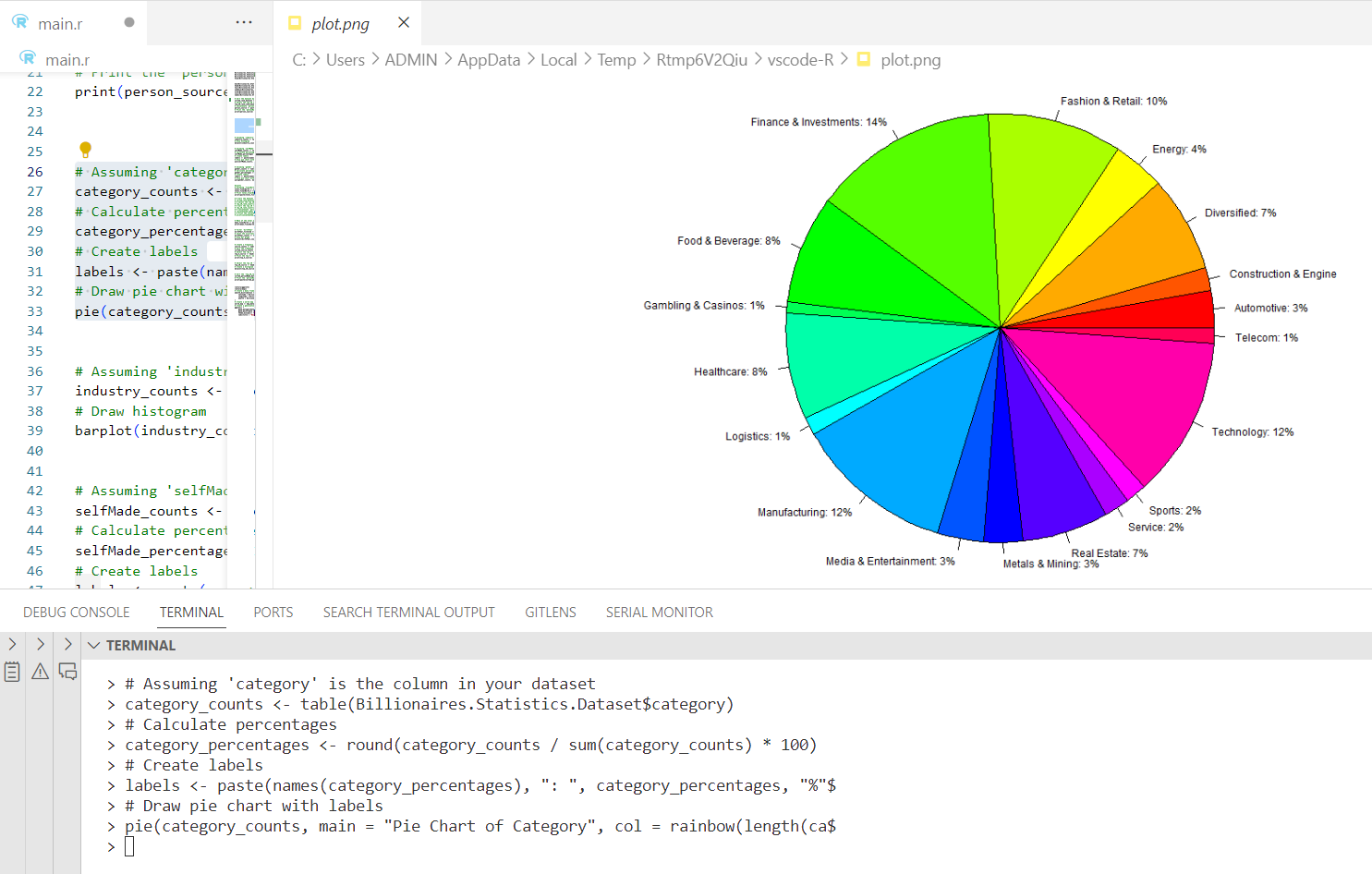




## Data visualization

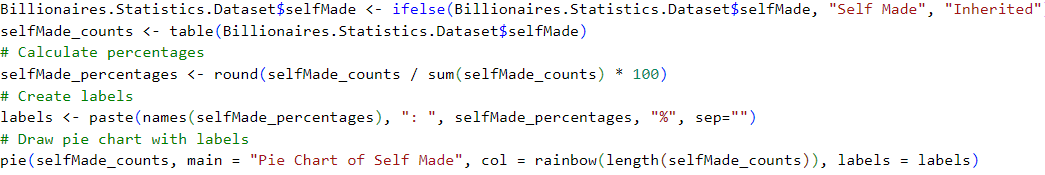
### Pie chart - 1

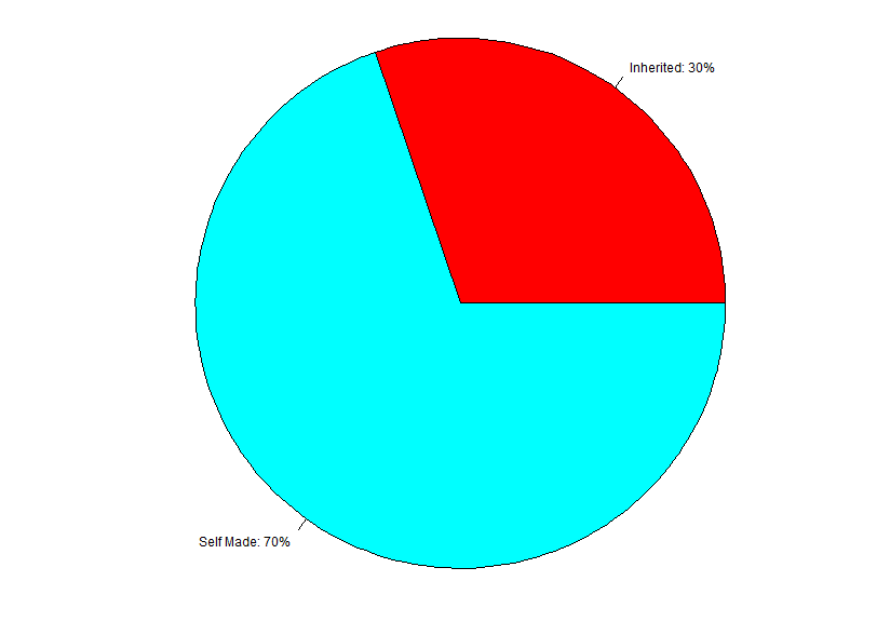
This code will give the percentage of sources from where the wealth is generated from category of work and produces a piechart accordingly.



### Pie chart - 2

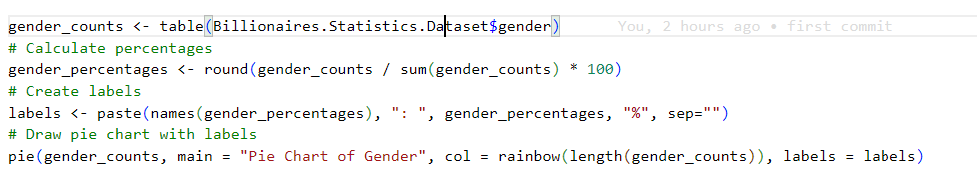
Calculating the ratio between self-made billionare and wealth inherited

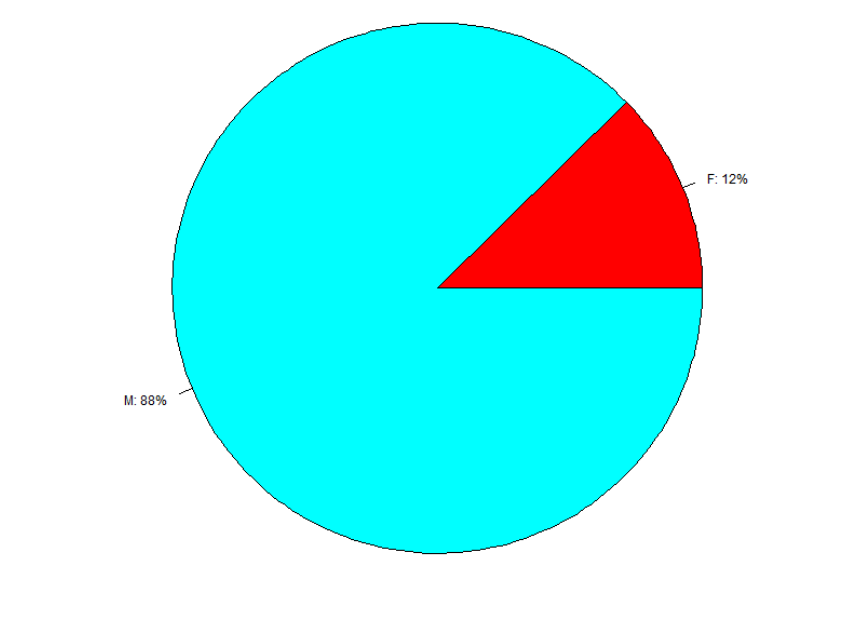




### Pie chart -3

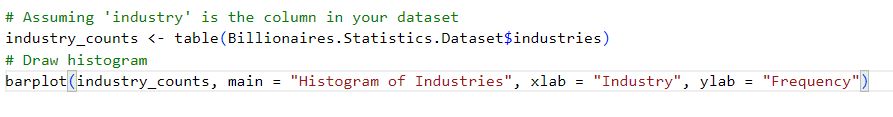
Ratio between male and female billionares

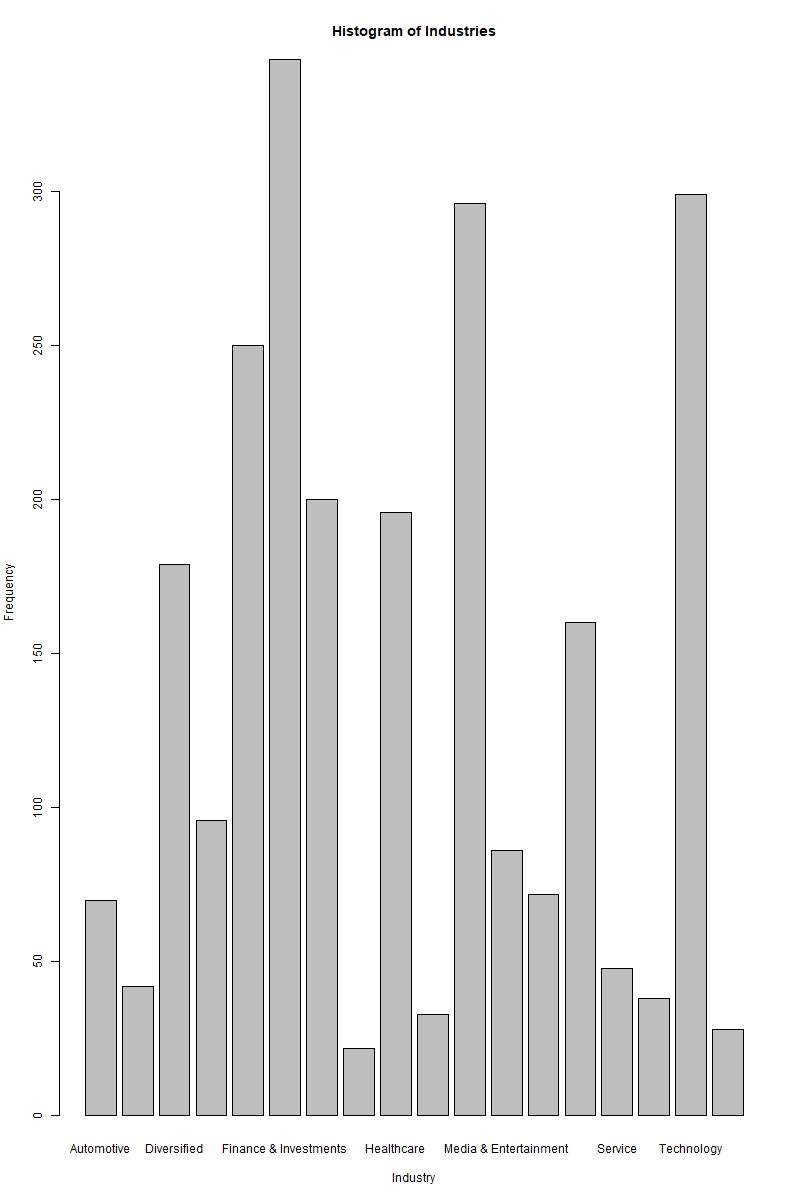




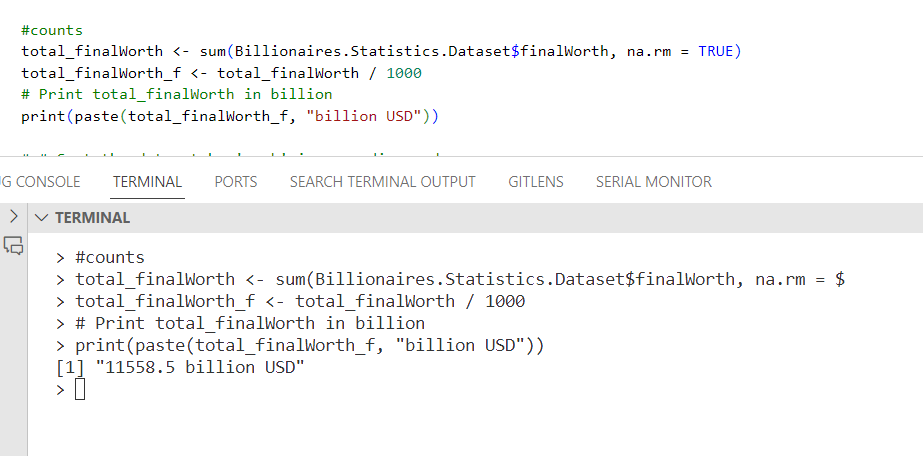
### Histogram

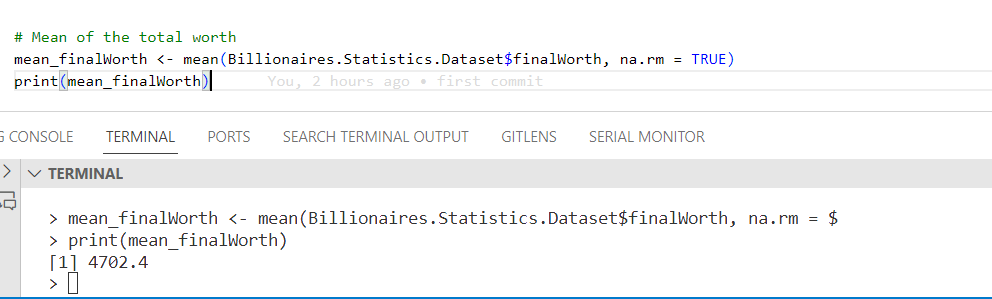
This code will give the percentage of sources from where the wealth is generated from industries and produces a bar graph accordingly.





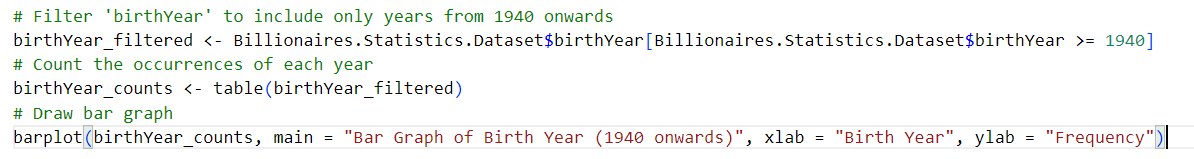
The total worth of all the billionares in the dataset.

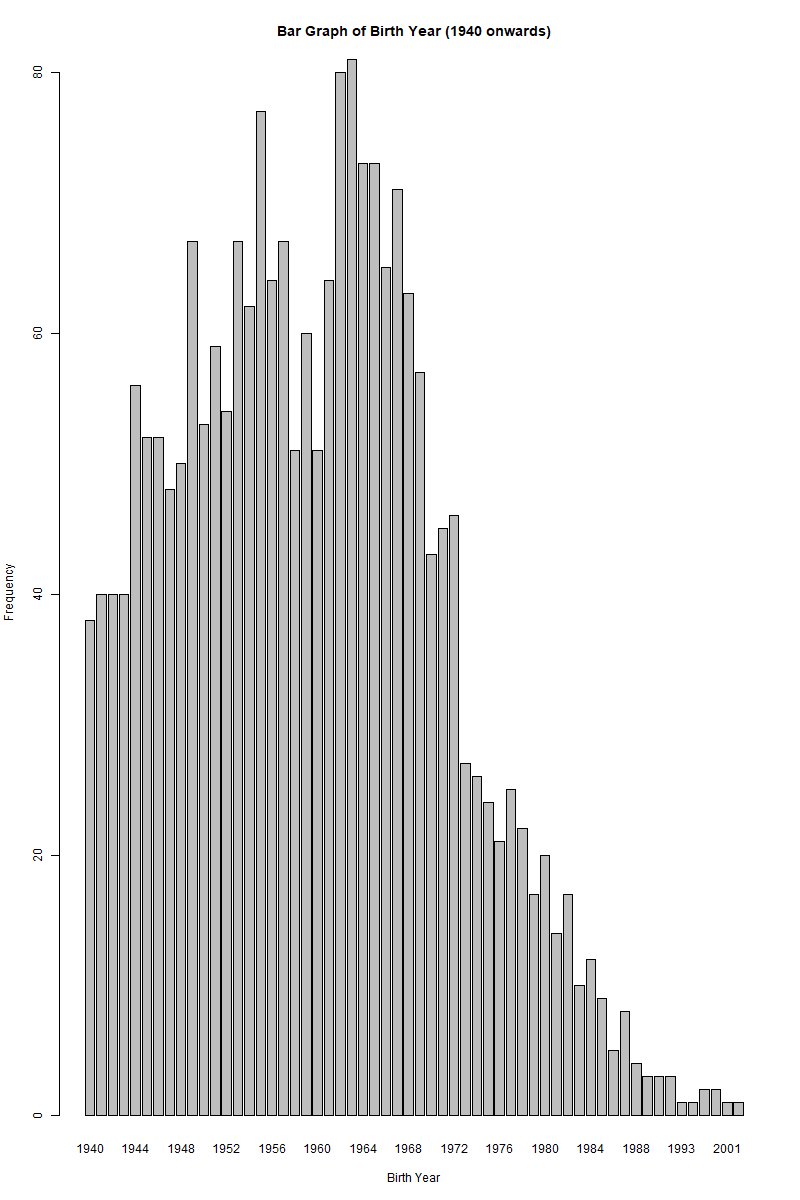


Mean of the worth.

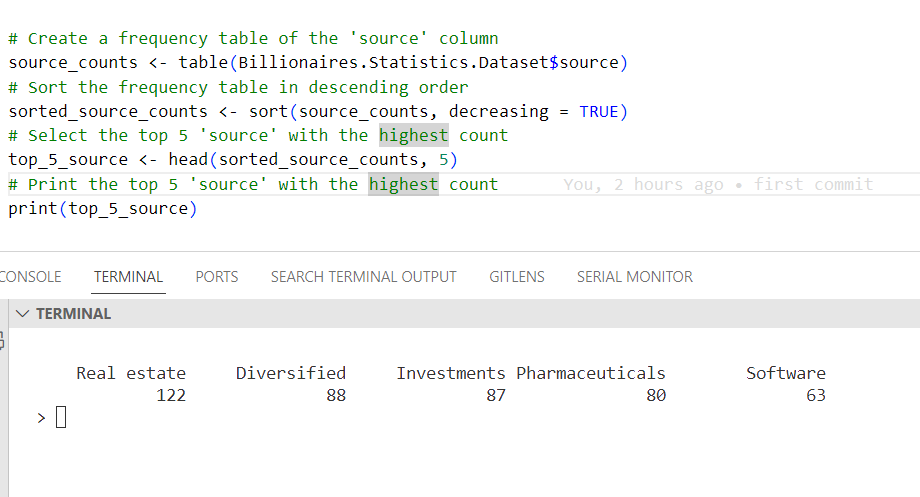
### Bar graph

A bar graph to represent the number of billionares born in the given year.

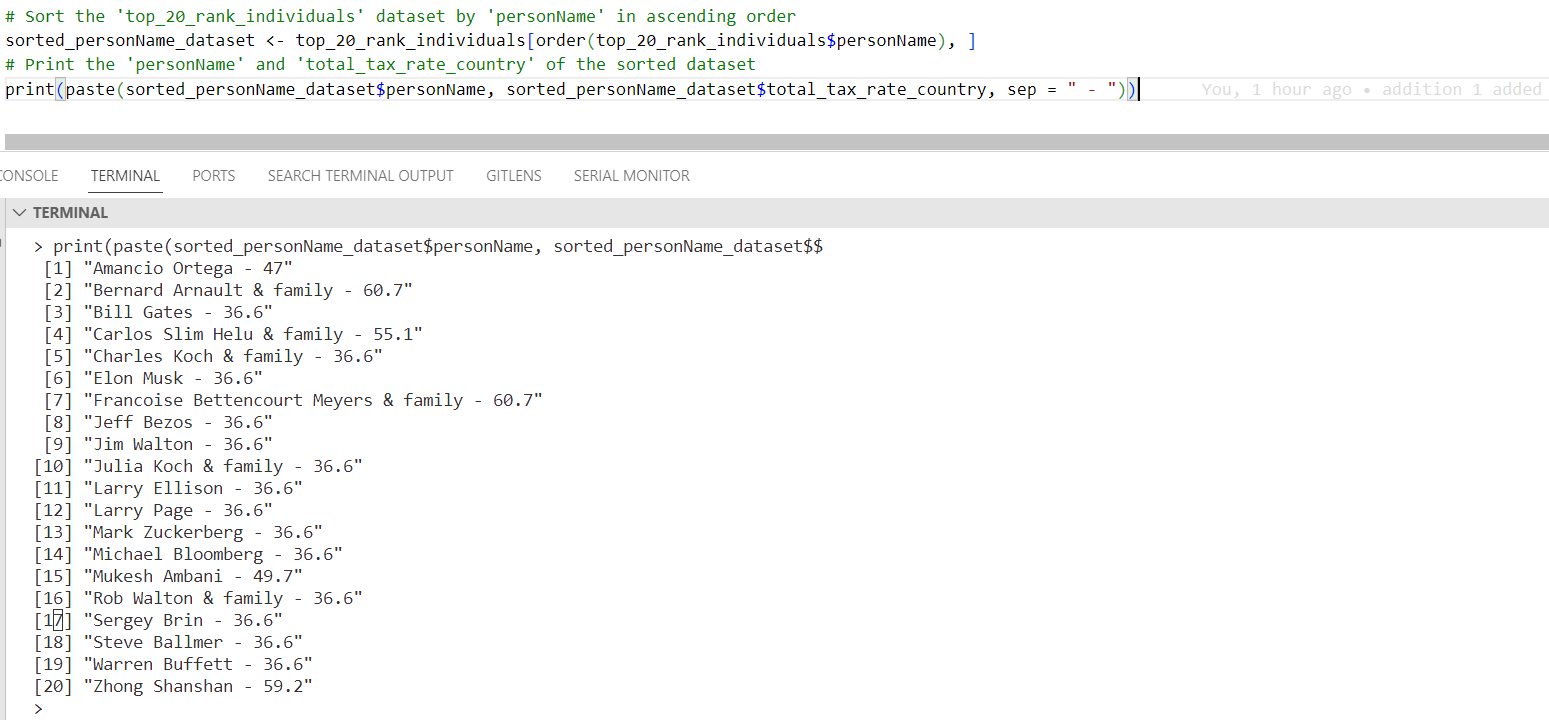




Billionares per sector



## Top 20 by tax rate



### Map graph

