

DATA ANALYSIS REPORT

OVERVIEW OF THE PROBLEM:

A survey has been prepared by two different agencies on different programming languages used in today's IT world and the data is recorded based on the use, rating, ranking and trending language characteristics. The problem is to perform a data analysis to extract hidden information from the data available. This process includes cleansing, transforming, visualizing and modelling data to discover useful information and to draw out informing conclusions.

This data analysis report will have based on the questions as follows:

1. Most trending programming language.
2. How different programming languages are correlated to each other?
3. Analysis based on the rank feature of different programming languages.
4. Which programming language have better and worst scope as per the market's share?
5. And some plots describing the data.

DATA DESCRIPTION:

Data recorded is in the format of an excel file with two different sheets. Each sheet contains data regarding the rank, rating and trend etc. from two different agencies. The data in two sheets are not exactly same but somehow supplement to each other.

Data in sheet1:

COLUMN NAMES	DATA TYPE
Rank	Numeric
Language	Character
Share	Character
Trend	Character

Data in sheet2:

COLUMN NAMES	DATA TYPE
Sep 2020	Numeric
Sep 2019	Numeric
Programming Language	Character
Ratings	Numeric
Change	Numeric

DIFFERENT STATISTICS:

Data in sheet1:

Language: 28 Programming Languages

Share: Top programming Languages with highest market share: PYTHON, JAVA, JAVASCRIPT, C#, PHP, C/C++, R

Trend: Python is the most trending language as per the first agency

Rank	Language	Share	Trend
Min. : 1.00	Length:28	Min. : 0.0192	Min. : -3.1000
1st Qu.: 7.75	Class :character	1st Qu.: 0.4975	1st Qu.: -0.1000
Median :14.50	Mode :character	Median : 1.0850	Median : 0.0000
Mean :14.50		Mean : 3.5028	Mean : -0.1662
3rd Qu.:21.25		3rd Qu.: 3.1125	3rd Qu.: 0.0020
Max. :28.00		Max. :31.5600	Max. : 0.0290

Data in sheet2:

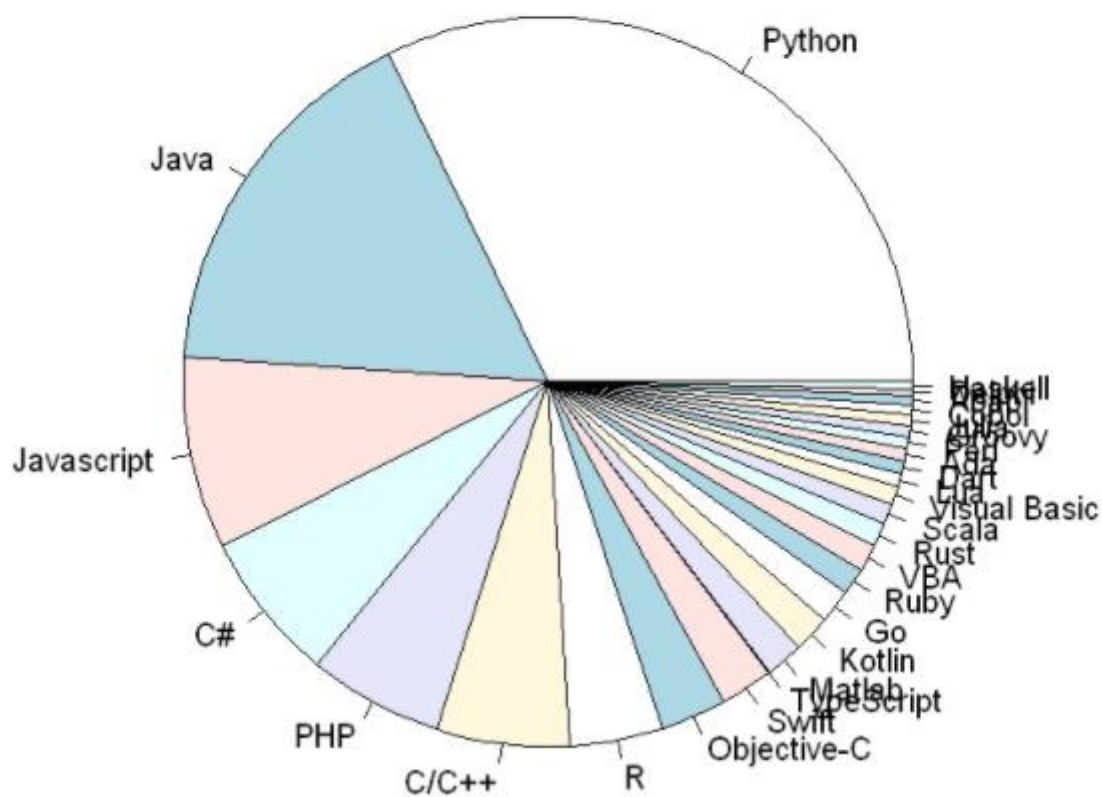
Language: 20 Programming Languages

Ratings: Top rated programming Languages: JAVA, C/C++, PYTHON, C#, JAVASCRIPT, VISUAL BASIC, R, PHP

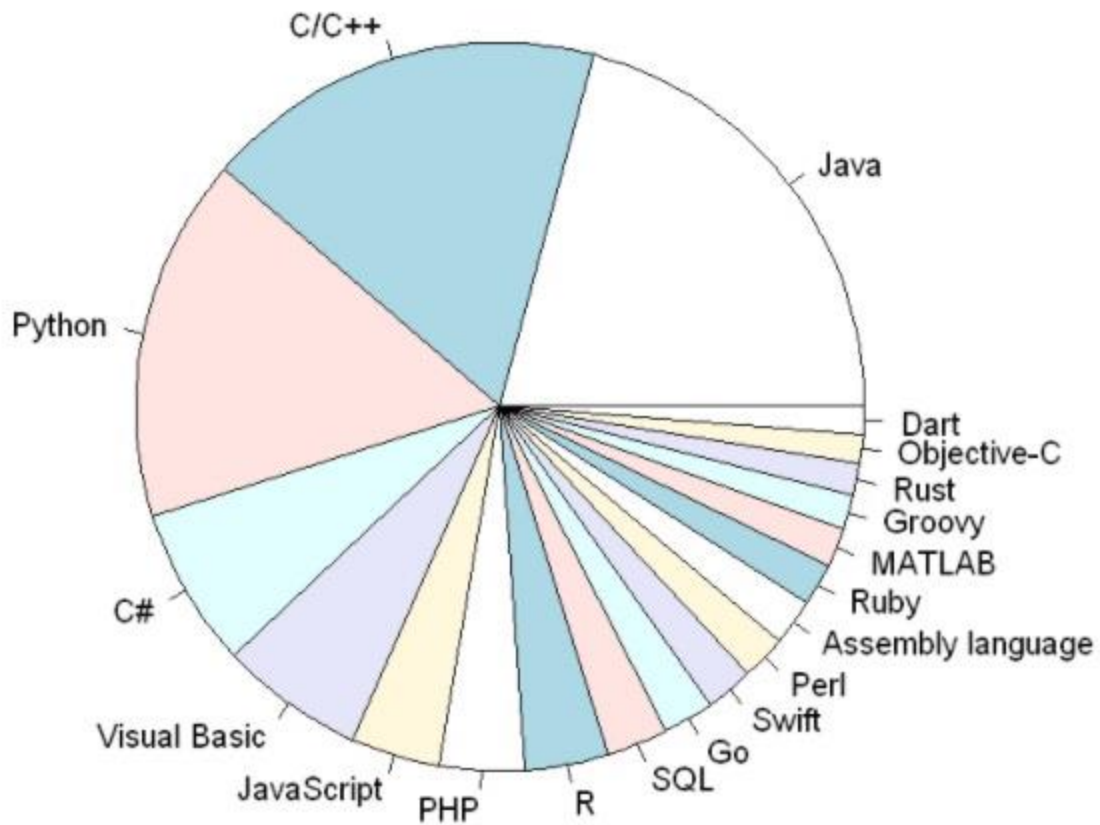
Rank_2020	Rank_2019	Programming Language	Ratings
Min. : 1.00	Min. : 1.00	Length:20	Min. : 0.00770
1st Qu.: 5.75	1st Qu.: 5.75	Class :character	1st Qu.: 0.01205
Median :10.50	Median :10.50	Mode :character	Median : 0.01610
Mean :10.50	Mean :11.85		Mean : 0.03809
3rd Qu.:15.25	3rd Qu.:16.50		3rd Qu.: 0.04235
Max. :20.00	Max. :33.00		Max. : 0.15950
Change			
Min. : -0.031800			
1st Qu.: 0.000025			
Median : 0.002700			
Mean : 0.001875			
3rd Qu.: 0.006500			
Max. : 0.014800			

PLOTS:

Pie chart for market share of languages



Pie chart for Rating of Programming languages



CONCLUSION:

The data analysis report provides the following overview:

- Python is the most trending language with the highest coverage of market share.
- Some programming languages are in the verge of being obsolete and has no change in the ratings for a considerable amount of time.
- R being the language in the list of top 10 high rated languages.