A Comparative Study of SAS, R, and Python in Business Intelligence

Soumita Banerjee
Dr. Illya Mowerman
Department of Technology Management

Introduction

In recent years, SAS, R, and Python are the three predominant analytical tools which are being used in small, medium, and large scale industries. SAS, R, and Python are not the only BI tools, but they are most popular programming languages that are used in BI. This study aims to analyze the market share of these three tools and their roles in Business Intelligence.

Research Methodology

- The related articles and journals were collected from UB digital library, Google Scholar, different blogs and websites.
- The report is review-centric.
- Review centric research process follows the three phases: (1) brainstorming; (2) narrowing down; (3) evaluation.

Business Intelligence

In today's world, BI is a must. It is a standard practice for all organization. BI is crucial for success in competitive business world. Most companies use BI in one way or another. BI is the combination of technological tools, systems, applications, and practices. It is a support system which stores and analyze historical business data to make accurate business decisions.

According to Mulcahy (2007), Business Intelligence is the term which refers to varieties of applications to

- Analyze raw data through data mining,
- Online analytical processing,
- Querying and Reporting.

R

- R was developed by 1995 by Ross Ihaka and Robert Gentleman in at the University of Auckland, New Zealand
- It is a dialect of S language, developed in Bell Labs and supported by the R Foundation for Statistical Computing
- Open source makes R language free & easily available
- Packages are quickly available for installing and updating
- R. It is useful for the following tasks:
  - Manipulating packages and strings
  - Analyze data with regular and irregular time series
  - Data visualization
  - Machine learning
- R is being used in large companies like Bank of America, Ford, Uber, Foresquare etc. as their primary analytical tool

Python

- Python was developed by Guido van Rossum in Netherland in late 1980s and was implemented in December 1989.
- The availability of Python as an open source software makes it easily available
- Python needs programming codes to analyze data
- Python is useful for:
  - Developing websites
  - Computing scientific and numeric data
  - Creating new software
- Python is being used in companies like Yahoo, NASA, IBM, Mozilla, Dropbox etc. as their primary data analysis tool.

Research Findings

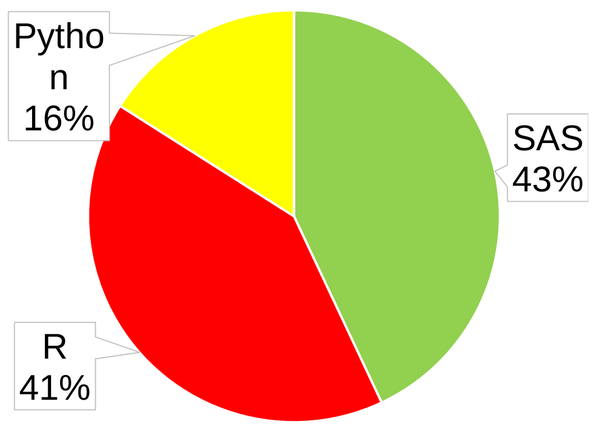

Figure 1. Tool Preference Distribution for Predictive Analysis

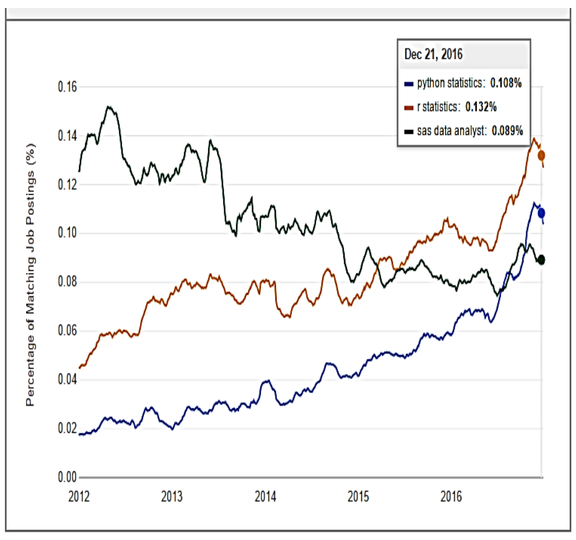

Figure 2. Job Trends of SAS, R & Python programmers

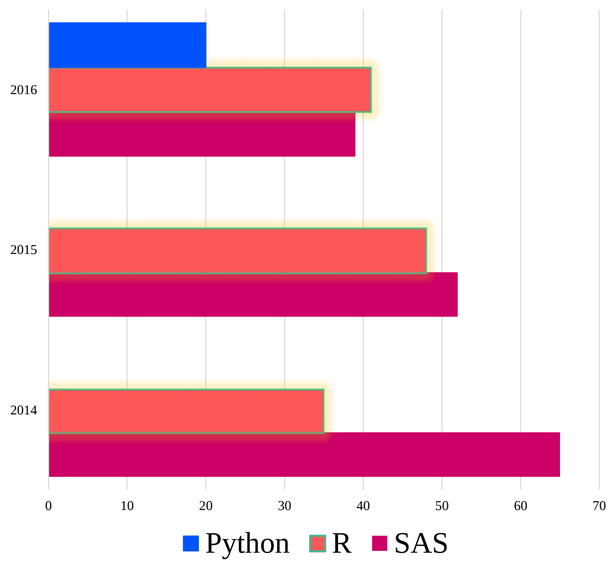

Figure 3. Tool Preference by Time

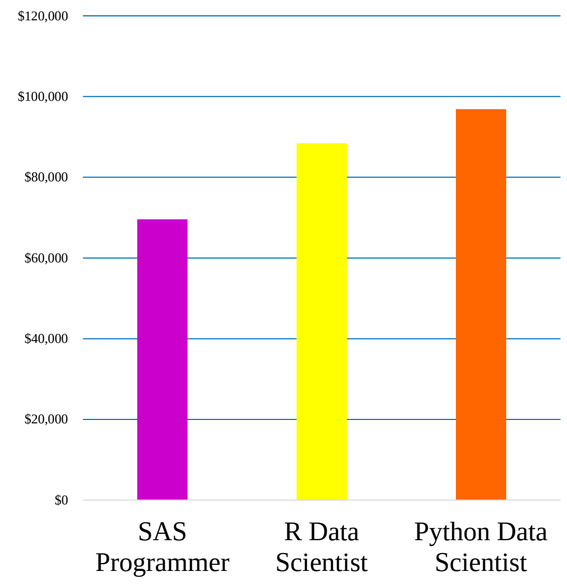

Figure 4. Salary Comparison

SAS

- It was started by James Howard Goodnight in July 1, 1976
- SAS is an integrated system of software which supports statistical functions by both ways – graphical user interface and programming language.
- SAS is useful for the following tasks:
  - Data storage, retrieval, and maintenance
  - Report generation and graphical presentation
  - Statistical and mathematical analysis
  - Predicting Business processes and decision supports
- Companies like Google, Facebook, Nestle, Volvo, Barclays, HSBC are using SAS as their primary data analytical tool.

Forecasting Trends

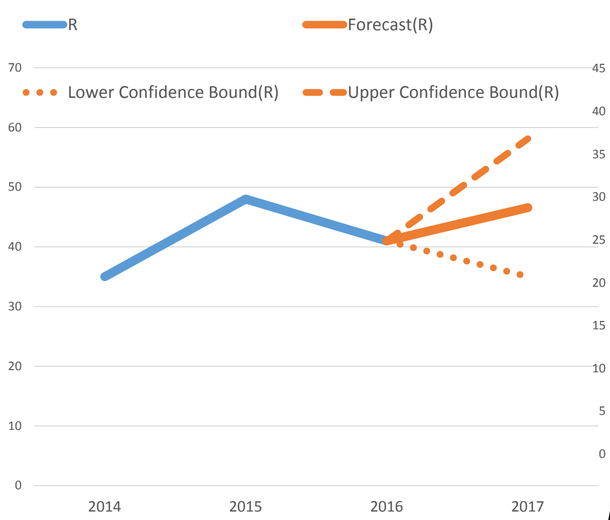

Figure 5. Predicting R Preference

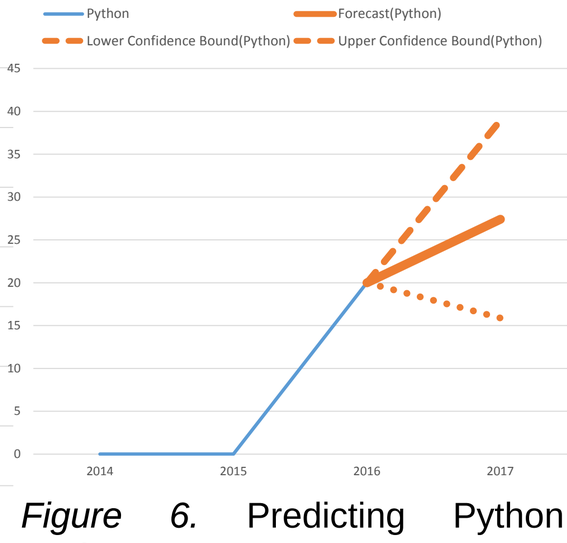

Figure 6. Predicting Python Preference

Conclusion

Although, none of them will disappear in near future, chances are that R and Python will become the most preferable BI tools in long term. This likely due to the adoption of open source software. However, historical data shows that preferences can change with new advances. The choice between these three tools depends on many factors which are both internal and external to organizations. The survival of all these tools are highly dependent on their continue to innovate.

Reference

•Average Salary for Skill: Python. (February 21, 2017). Retrieved February 26, 2017, from http://www.payscale.com/research/US/Skill=Python/Salary

•Average Salary for Skill: R. (February 20, 2017). Retrieved February 26, 2017, from http://www.payscale.com/research/US/Skill=R/Salary

•Burtch, Works. (July 13th, 2016, February 25). SAS, R, or Python Survey 2016: Which Tool Do Analytics Pros Prefer? Retrieved from http://www.burtchworks.com/2016/07/13/sas-r-python-survey-2016-tool-analytics-pros-prefer/

•SAS Programmer Salaries in US. (February 25, 2017). Retrieved February 26, 2017, 2017, from https://www.indeed.com/salaries/SAS+Programmer-Salaries,-United+States