International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | Jun 2022 www.irjet.net p-ISSN: 2395-0072

Comparing the performance of a business process: using Excel & Python

Pranav Chaudhari, Sunny Nahar

Department of Masters of computer Applications,
Vivekanand Education Society's Institute of Technology, Chembur, Mumbai 400074, India

---

Abstract - *In recent times due to technological advances the need to improve the existing system of processes is greater than ever before. Also, with data-driven decision-making processes that come into play the need to integrate that with the process helps business and management in making the right decision that benefits everyone. This paper focuses on automating existing processes using Python as a translation method, understanding data thus helping decision-driven Data by highlighting issues and providing details in data analysis.*

Key Words*: *Data Analysis, Business Analysis, Automation, Python, Big Data, Machine Learning, Artificial Intelligence

1. INTRODUCTION

1.1 Data Analysis

Data analysis is the process of evaluating, cleaning, modifying, and modeling data for the purpose of obtaining useful information, informing conclusions, and supporting decision-making. Data analysis has many features and methods, which include different strategies under different terms, and are used in different domains of business, science, and social sciences. In today's business world, data analysis plays a key role in making the decisions more scientific and helps businesses more efficient. [4]

1.2 Business Analysis

Business Statistics (BA) refers to the skills, expertise, and processes of repeated evaluation and past business performance research to gain an understanding and advance business planning. Business statistics focus on developing new data and understanding business performance based on data and statistical methods. In contrast, business acumen is traditionally focused on using a consistent set of metrics to both measure previous performance and direct business planning. In other words, business intelligence focuses on definition, whereas business analysis focuses on physician prediction and writing. [3]

1.3 Python for Automation

The most important skills for business analysts in consultation with executives, bank investments, and many other analytics activities (usually at least traditionally) were Excel and PowerPoint. many have become very visible in recent years. The data and its complexity are growing exponentially, requiring advanced data processing tools such as programming languages. In this article, the emphasis is on Python and how it works, and why we believe it is an important skill right now for the future. [1][2]

2. LITERATURE SURVEY

Nikita Khudov [1] describes few characteristics why one should opt for automation using python for performing the business analysis. The idea is to demonstrate certain metrics that gives python the upper hand for better decision making and time saving techniques.

sayoneadmin [2] describes the business analytics trends that are go hand in hand with python to give us the best fit to help us analyze the data and make decisions out of it.

Business analysis [3] is the expert advice of identifying business needs and finding solutions to business problems. Solutions often involve part of software development, but may also include process development, organizational change or strategic planning and policy development.

Data analysis [4] has many features and methods, which include different strategies under different names, and are used in different domains of business. There are various processes within it ranging from data requirements, to its cleaning, testing and modeling.

3. PROPOSED SYSTEM

The proposed system will help users to upload, download data using python as a tool, analyzing data to provide information that can help the business team make decisions to benefit the most from it. We can also store data in different repositories and create different dashboards to present data in a more informative way. Ultimately it will help save a lot of time, effort and risk of human error in order to provide solutions.

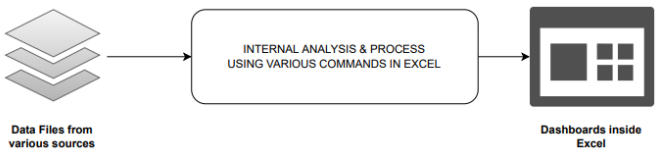

Fig -1: Manual Process

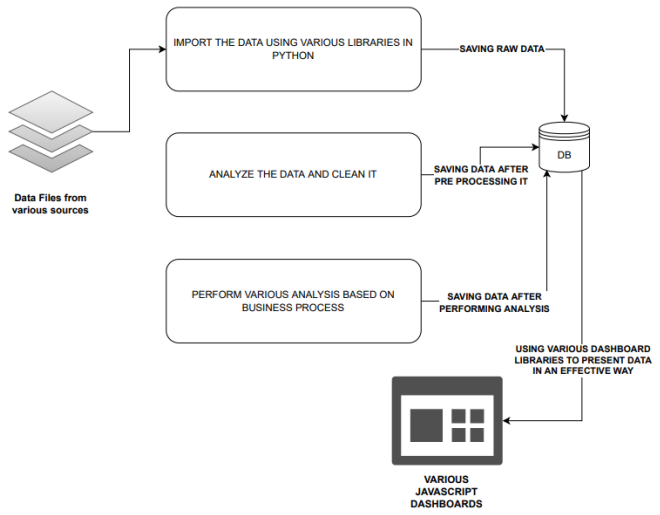

Fig -2: Proposed Process

4. METHODOLOGY

While testing the performance on data analytics using tools such as excel and python, I tested the functionality in the following parameters:

4.1 The Amount of data that can be handled

We have used different sets of data with different values to test the performance of both excel and python. In small databases both are advanced, and python works very fast and does not slow down. But when the no. of lines goes over 1 lac line objects is when the excel starts to loosen the smoothness and firmness of the software comes into play. We can also add a database at the end of the python code to save data for future use and usage.

4.2 Doing Analysis

In the second step, we used a few data sets to perform the analysis based on a few steps. It can be done in both Excel and python. But with excel the user should have some knowledge of the UI and where each function resides and can later filter data by using shortcut keys to filter data. But with python this process is much easier because the coding language is easier to understand, and the implementation is easier.

Following functionalities were tested in both excel and python:

A. Loading the data from different sources

1. In excel, we must repeat the import steps for each new data source needed for analysis, which over time does not provide easy import of data from multiple sources (this can be achieved using Power Query in Excel)
2. Whereas in python, it only one step that can be inserted inside the code with the location of the new source and then the code can simultaneously import data from multiple sources

B. Filtering data on certain criteria

1. In excel, we can filter data on various terms, but it takes time to add terms and filter data from time to time
2. Whereas in python, we can add many conditions and convert data in a matter of seconds and then compare data using data frames in python

C. Finding duplicates from the dataset based on specific criteria

1. In excel, it is possible to find duplicates, but it takes 2-3 different steps based on the concept to identify duplicate data and filter that in the final data.
2. Even in python, a single line of code can help detect and remove duplicates. In this case we can say the diagnostic parameter, which data we should keep and thus remove other line / lines of duplicate data.

D. Changing the data format of individual cell or column to a particular format

1. In excel, it is sometimes very difficult to change the cell format of the value you want. Also, if the data is left unchanged then it may affect the analysis
2. Whereas in python, after the process flow is set and encoded. Conversion will happen automatically, and the conversion type is also very fast thus speeding up the whole process

E. Merging data from two or more datasets at a same time

1. In excel, we use lookup functions to combine data from two different sets based on common parameters. Post that we need to make changes by saving the lookup data as values to create additional lookups
2. While in python, we just need to specify a join type (like SQL query) and python will do everything else. The performance while performing joins - the speed it is very fast even on large databases

4.3 Scalability & Future use

In more recent times, with new technologies being added to everyday processes, using python not only provides an easy-to-use method but also helps to balance performance based on changing trends. With the support of huge community and new libraries added like ML, AI we can use that to analyze data and predict output. Since Big Data is already in use, we can use ML / AI to process historical data and predict a business-benefit outcome.

4.4 Dashboards & Displaying outputs

Recent trend dashboards are used to display output based on analytics. Dashboards can be created on both excel and python. An added benefit of creating a dashboard in python is that we can use various JavaScript libraries to present the output in a highly informative way. We just need to link the input data with the correct ID and the dashboards are automatically created in the front part of the web app.

4.5 Data Integrity & Safety

Another advantage of using python is keeping data (raw + analyzed) in a secure database for future use. When we use this practice, we are able to maintain the integrity of the data and avoid the risk of violating the privacy of our data by not sharing data with anyone not involved in the process. With excel it is possible that the data may be corrupted, or the data may fall into the wrong hands that can all be avoided by using python and database in an end-to-end process.

4.6 Easy of Access & Real Time Sharing

With the concept of centralized data and real-time sharing, data is updated simultaneously as webapp, and python will be connected to the same database. It is also easy to share as users can be anywhere in the world and a mapping is created while a project is created in a web application.

5. RESULT

Based on the analysis and evaluation of end-to-end performance, we have come to understand that the use of python as an automated method of creating webapps is superior, helpful and in recent times very useful to drive the whole process, it also uses data more efficiently and provide information based on current trends and business needs.

REFERENCES

[1] https://towardsdatascience.com/why-python-is-essential-for-business-analysts-ed3d5a2b194c

[2] https://www.sayonetech.com/blog/why-python-important-business-analytics/

[3] https://en.wikipedia.org/wiki/Business_analysis

[4] https://en.wikipedia.org/wiki/Data_analysis