**A Project Report**

**On**

***Business Sales Data Analysis***

Submitted by:

Vishnu Vyas, Rishabh Jain, Rishit Bajaj

Under the guidance of

**Mr. Prabhat Das**

(PGT-IP)



**ARMY PUBLIC SCHOOL JORHAT**

**MES GATE ROWRIAH JORHAT**

**Assam-785004**

**CERTIFICATE BY PRINCIPAL**

This is to certify that this project report entitled “**Business Sales Data Analysis**” submitted by **Vishnu Vyas, Rishabh Jain, Rishit Bajaj** to Army Public School Jorhat has been examined and evaluated.

The report has been prepared as per the regulations of CBSE and qualifies to be accepted.

Date:

Place:

Mrs. Firdausi Sultana Hazarika

(Principal)

Army Public School Jorhat

**CERTIFICATE BY EXAMINERS**

This is to certify that this project report entitled “**Business Sales Data Analysis**” is the bona fide work of who carried out the project work under my supervision and guidance.

To the best of my knowledge, the matter embodied in the report has not been submitted to any other institute for the award of any other degree.

Date:

Place:

Mr. Prabhat Das

(External Examiner) (Internal Examiner)

**ACKNOWLEDGEMENT**

I take this opportunity to extend my heart full gratitude to Army Public School Jorhat for providing me the opportunity.

I am highly grateful to my guide Mr. Prabhat Das, PGT-IP, Army Public School Jorhat for giving us the opportunity to work under him and providing us an ample guidance and support through the project.

Lastly, I would also like to thank the authors whose publications guided us regarding our project.

**DECLARATION**

I admit that this report is of my own work and all the sources of the information used in this report have fully acknowledged.

I hereby declare that the dissertation work entitled “**Business Sales Data Analysis**” submitted to the Army Public School Jorhat, is prepared by me and was not submitted to any other institution for award of any other degree.

Date:

Place:

Signature

**Abstract**

In the era of today’s world, where the globe is increasing it’s fecundity monumentaly, the introduction of computers has contributed to a boundless scale to this everlasting race. The invention of computers and mechanisation has eased the human difficulties, where one had to store all the data in the form of physical means, adding to the efforts, space and expenditure.

With the introduction of databases, one can store a colossal data in the form of bytes and bits in a single computer, compared to the earlier times where one had to keep a raw data in the form of files, folders, cupboards, etc., which had to get themselves put in a huge amount of effort. All the business giants in today’s era have the use of databases as their primary mode of storage.

The present program lays out the user the “Business Sales Data Analysis” of the prominent Amazon inc., which is one of the foremost companies of the globe, known for it’s online stores and services. The user here inputs their e-mail, which then receives an OTP through the program on the mail itself’ asking the user to authenticate the code. On the successful authentication if the code , the user is provided with the list of options, showing them a list of options from the database created in the program. The user is to select the option(s) they wish to see from the given list by feeding in the numbers allotted to them. On feeding in the value(s), they are displayed with the information stored in the database in the form of figures of graphs.

**Contents**

* Tools and Library 1
* Introduction 3
* Project overview 4
* Source code 11
* Commands in SQL 16
* Conclusion and Future work 17
* References 18  **List Of Figures**
* Fig 1.
* Fig 2.
* Fig 3.
* Fig 4.
* Fig 5.
* Fig 6.
* Fig 7.
* Fig 8.
* Fig 9.
* Fig 10.
* Fig 11.
* Fig 12.
* Fig 13.

**Tools and Libraries Used**

**MySQL**

MySQL is the world's most used relational [database management system](https://www.limswiki.org/index.php?title=Database_management_system&action=edit&redlink=1) (RDBMS) that runs as a server providing multi-user access to a number of databases. It is named after developer Michael Widenius' daughter, My. The phrase "SQL" stands for "structured query language."

[Open source projects](https://www.limswiki.org/index.php/Category:Open-source_software) that require a full-featured database management system often use MySQL. Applications which use MySQL databases include:

[TYPO3](https://www.limswiki.org/index.php/TYPO3), [Joomla](https://www.limswiki.org/index.php/Joomla), [WordPress](https://www.limswiki.org/index.php/WordPress), [phpBB](https://www.limswiki.org/index.php/PhpBB), [Drupal](https://www.limswiki.org/index.php/Drupal), and other software built on the LAMP software stack. MySQL is also used in many high-profile, large-scale Web products, including Wikipedia, Google (though not for searches), Facebook, and Twitter.

The free open-source version of MySQL is commonly referred to as the MySQL Community Edition. For commercial use, several paid editions are also available, offering additional functionality.[1]

**PyCharm**

PyCharm is an [integrated development environment](https://en.wikipedia.org/wiki/Integrated_development_environment) (IDE) used in [computer programming](https://en.wikipedia.org/wiki/Computer_programming), specifically for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) language. It is developed by the [Czech](https://en.wikipedia.org/wiki/Czech_Republic) company [JetBrains](https://en.wikipedia.org/wiki/JetBrains).[[5]](https://en.wikipedia.org/wiki/PyCharm#cite_note-5) It provides code analysis, a graphical debugger, an integrated unit tester, integration with [version control systems](https://en.wikipedia.org/wiki/Revision_control) (VCSes), and supports web development with [Django](https://en.wikipedia.org/wiki/Django_(web_framework)) as well as [data science](https://en.wikipedia.org/wiki/Data_science) with [Anaconda](https://en.wikipedia.org/wiki/Anaconda_(Python_distribution)).

PyCharm is [cross-platform](https://en.wikipedia.org/wiki/Cross-platform), with [Windows](https://en.wikipedia.org/wiki/Windows), [macOS](https://en.wikipedia.org/wiki/MacOS) and [Linux](https://en.wikipedia.org/wiki/Linux) versions. The Community Edition is released under the [Apache License](https://en.wikipedia.org/wiki/Apache_License), and there is also Professional Edition with extra features – released under a [proprietary license](https://en.wikipedia.org/wiki/Proprietary_software). [2]

**Pandas**

In [computer programming](https://en.wikipedia.org/wiki/Computer_programming), pandas is a [software library](https://en.wikipedia.org/wiki/Software_library) written for the [Python programming language](https://en.wikipedia.org/wiki/Python_(programming_language)) for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and [time series](https://en.wikipedia.org/wiki/Time_series). It is [free software](https://en.wikipedia.org/wiki/Free_software) released under the [three-clause BSD license](https://en.wikipedia.org/wiki/3-clause_BSD_license). The name is derived from the term "[panel data](https://en.wikipedia.org/wiki/Panel_data)", an [econometrics](https://en.wikipedia.org/wiki/Econometrics) term for data sets that include observations over multiple time periods for the same individuals. Its name is a play on the phrase "Python data analysis" itself. [Wes McKinney](https://en.wikipedia.org/wiki/Wes_McKinney) started building what would become pandas at [AQR Capital](https://en.wikipedia.org/wiki/AQR_Capital) while he was a researcher there from 2007 to 2010. [3]

**Numpy**

NumPy (pronounced [/ˈnʌmpaɪ/](https://en.wikipedia.org/wiki/Help:IPA/English) ([NUM-py](https://en.wikipedia.org/wiki/Help:Pronunciation_respelling_key)) or sometimes [/ˈnʌmpi/](https://en.wikipedia.org/wiki/Help:IPA/English) ([NUM-pee](https://en.wikipedia.org/wiki/Help:Pronunciation_respelling_key))) is a library for the [Python programming language](https://en.wikipedia.org/wiki/Python_(programming_language)), adding support for large, multi-dimensional [arrays](https://en.wikipedia.org/wiki/Array_data_structure) and [matrices](https://en.wikipedia.org/wiki/Matrix_(math)), along with a large collection of [high-level](https://en.wikipedia.org/wiki/High-level_programming_language) [mathematical](https://en.wikipedia.org/wiki/Mathematics) [functions](https://en.wikipedia.org/wiki/Function_(mathematics)) to operate on these arrays. The ancestor of NumPy, Numeric, was originally created by [Jim Hugunin](https://en.wikipedia.org/wiki/Jim_Hugunin) with contributions from several other developers. In 2005, [Travis Oliphant](https://en.wikipedia.org/wiki/Travis_Oliphant) created NumPy by incorporating features of the competing Numarray into Numeric, with extensive modifications. NumPy is [open-source software](https://en.wikipedia.org/wiki/Open-source_software) and has many contributors. [[4]](https://en.wikipedia.org/wiki/NumPy#cite_note-Nature-5)

**Matplotlib**

Matplotlib is a [plotting](https://en.wikipedia.org/wiki/Plotter) [library](https://en.wikipedia.org/wiki/Library_(computer_science)) for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) programming language and its numerical mathematics extension [NumPy](https://en.wikipedia.org/wiki/NumPy). It provides an [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming) [API](https://en.wikipedia.org/wiki/API) for embedding plots into applications using general-purpose [GUI toolkits](https://en.wikipedia.org/wiki/GUI_toolkit) like [Tkinter](https://en.wikipedia.org/wiki/Tkinter), [wxPython](https://en.wikipedia.org/wiki/WxPython), [Qt](https://en.wikipedia.org/wiki/Qt_(software)), or [GTK+](https://en.wikipedia.org/wiki/GTK%2B). There is also a [procedural](https://en.wikipedia.org/wiki/Procedural_programming) "pylab" interface based on a [state machine](https://en.wikipedia.org/wiki/State_machine) (like [OpenGL](https://en.wikipedia.org/wiki/OpenGL)), designed to closely resemble that of [MATLAB](https://en.wikipedia.org/wiki/MATLAB), though its use is discouraged. [SciPy](https://en.wikipedia.org/wiki/SciPy) makes use of Matplotlib.

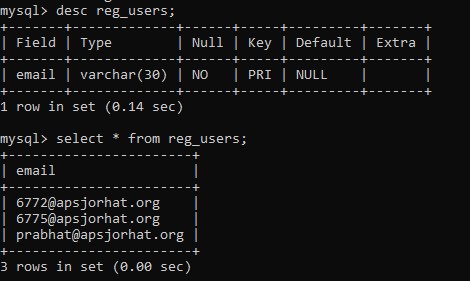
Matplotlib was originally written by [John D. Hunter](https://en.wikipedia.org/wiki/John_D._Hunter). Since then it has an active development community and is distributed under a [BSD-style license](https://en.wikipedia.org/wiki/BSD_licenses). Michael Droettboom was nominated as matplotlib's lead developer shortly before John Hunter's death in August 2012 and was further joined by Thomas Caswell.[[5]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-6)

**Introduction**

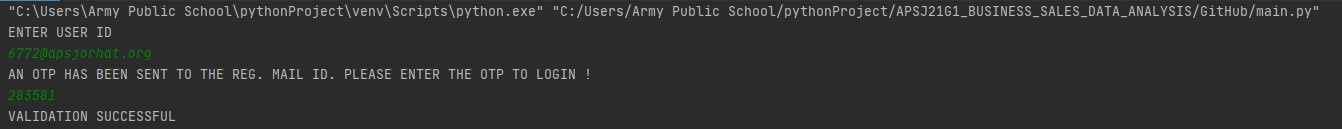
The theme of our project is ‘**BUSINESS SALES DATA ANALYSIS’**. This project is fine thought to make complex procedure of covid analysis in an easy manner which is systematic, modular designed, selective menu based user display. The modular design and constructed is very much user oriented in which user can easily understand the tools and can do edit of his own choice. The system is not any though more and does not possesses many application but it is made by focusing on maintaining record employee’s action in a computerized rather than time taking and cumbersome manual system.

The project software application that can easily handle by minimum educated and simple computer knowledge person without any option of error

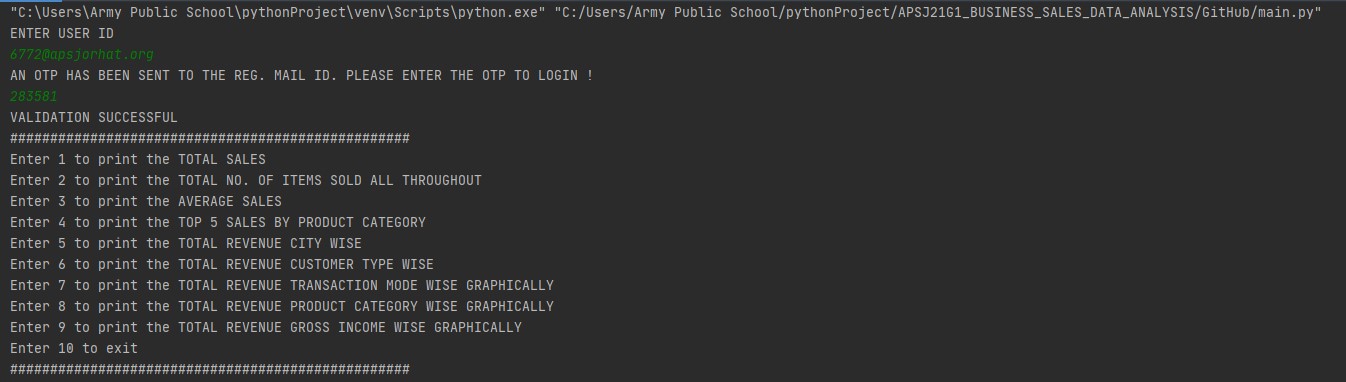
**Project Overview**



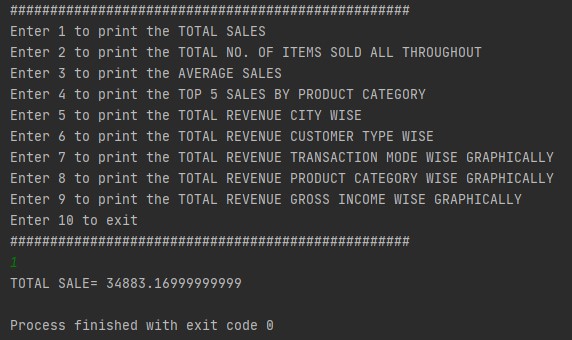
**Figure 1: All the information regarding the users are stored in this table**



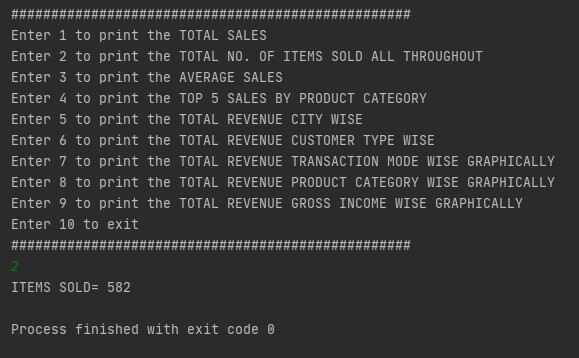
**Figure 2: When the home page is opened, program asks for user’s e-mail address, on entering an e-mail ID, program sends an OTP configuration to user’s e-mail**



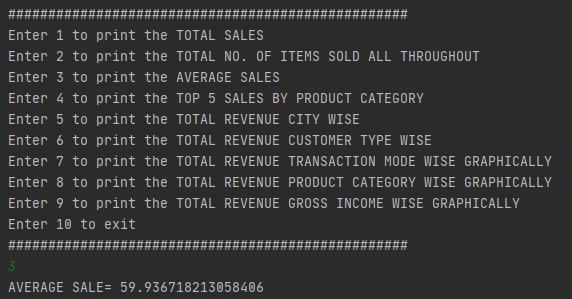
**Figure 3: After the login procedure is completed, the user is presented with the following list of options.**



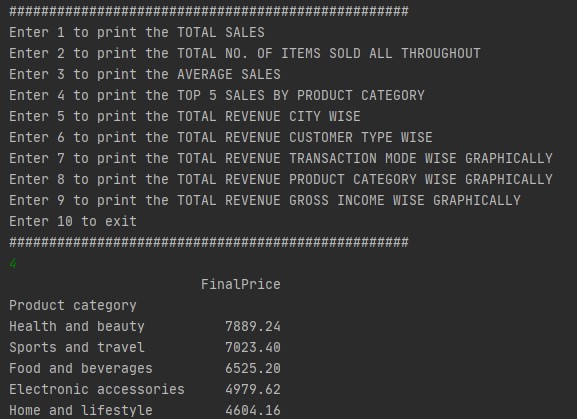
**Figure 4: Output for option 1(Total Sales)**



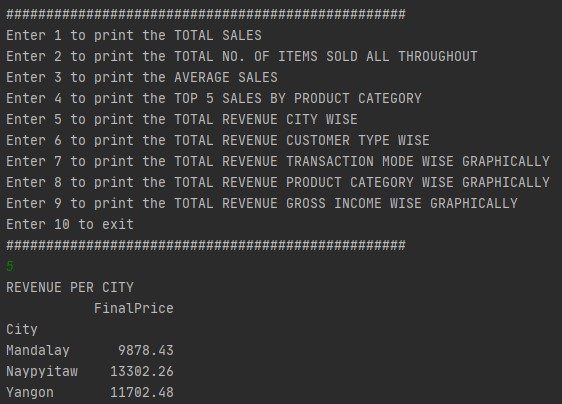
**Figure 5: Output for option 2 (Total no. Of Items Sold All Throughout)**



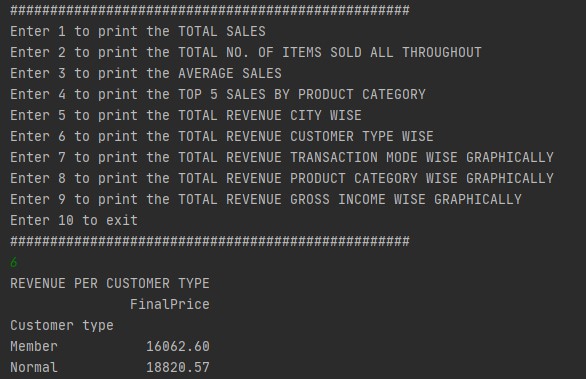
**Figure 6: Output for option 3 (Average Sales)**



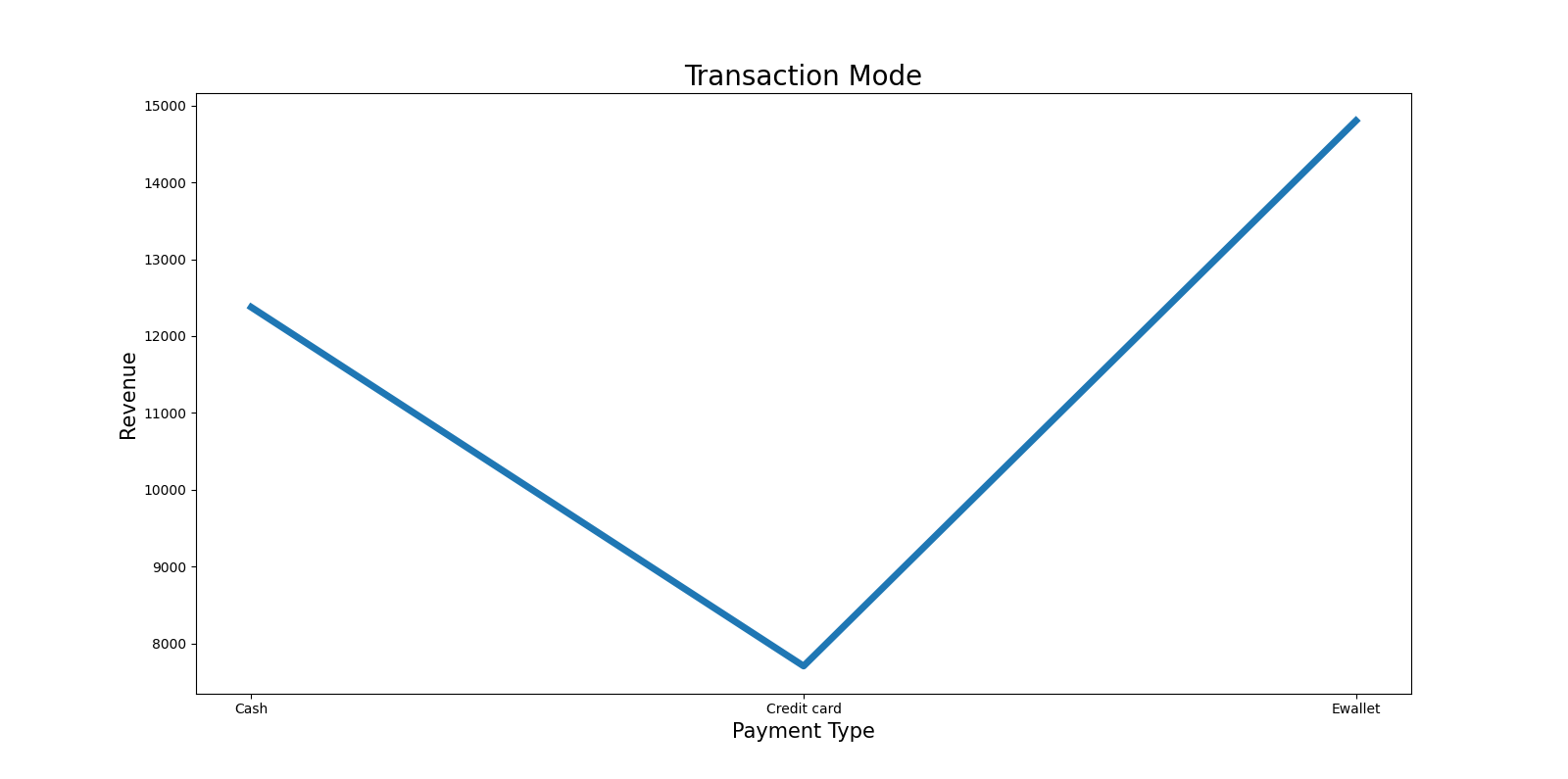
**Figure 7: Output for option 4(Top 5 Sales By Product Category)**



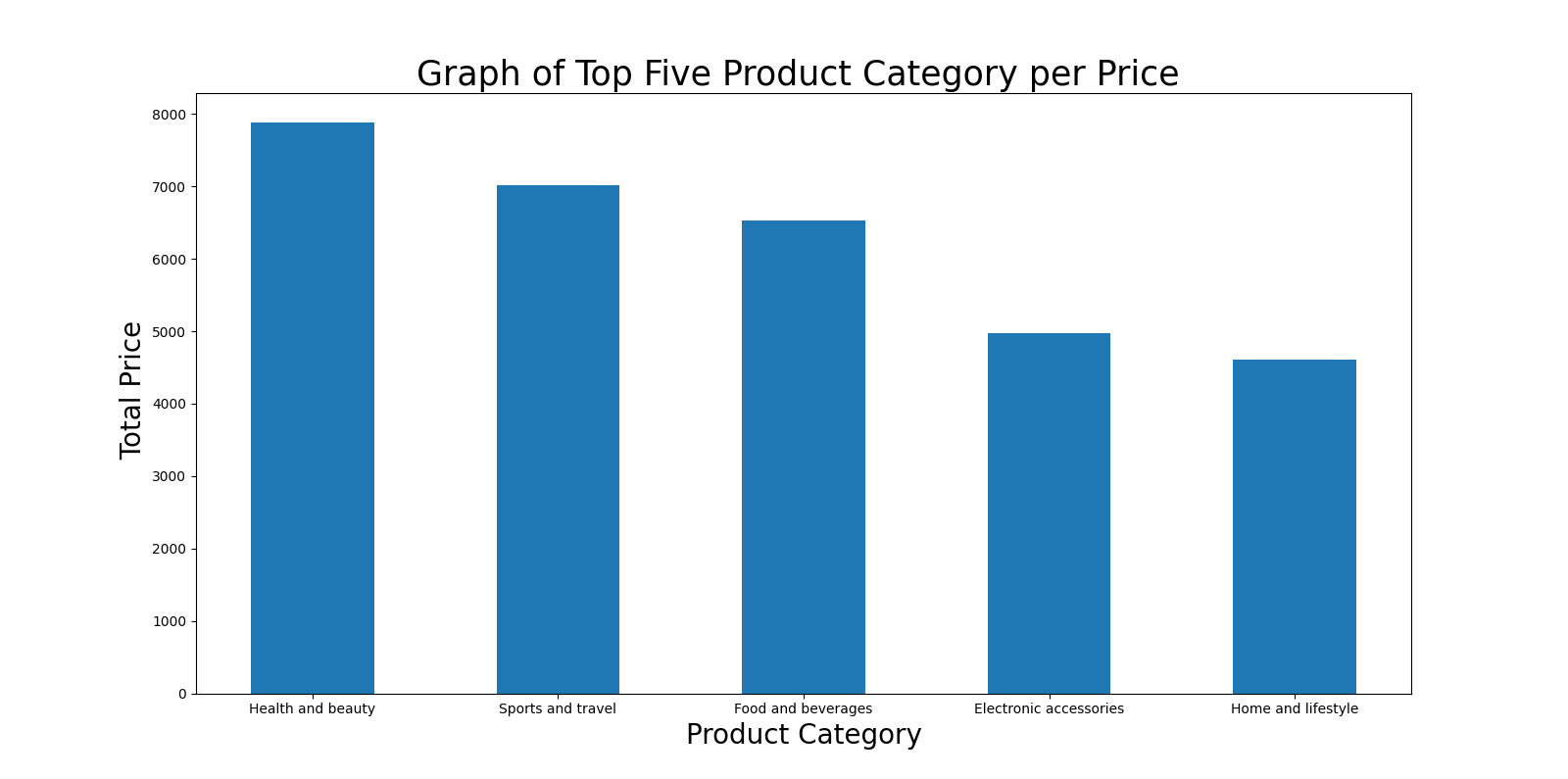
**Figure 8: Output for option 5(Total Revenue City Wise)**



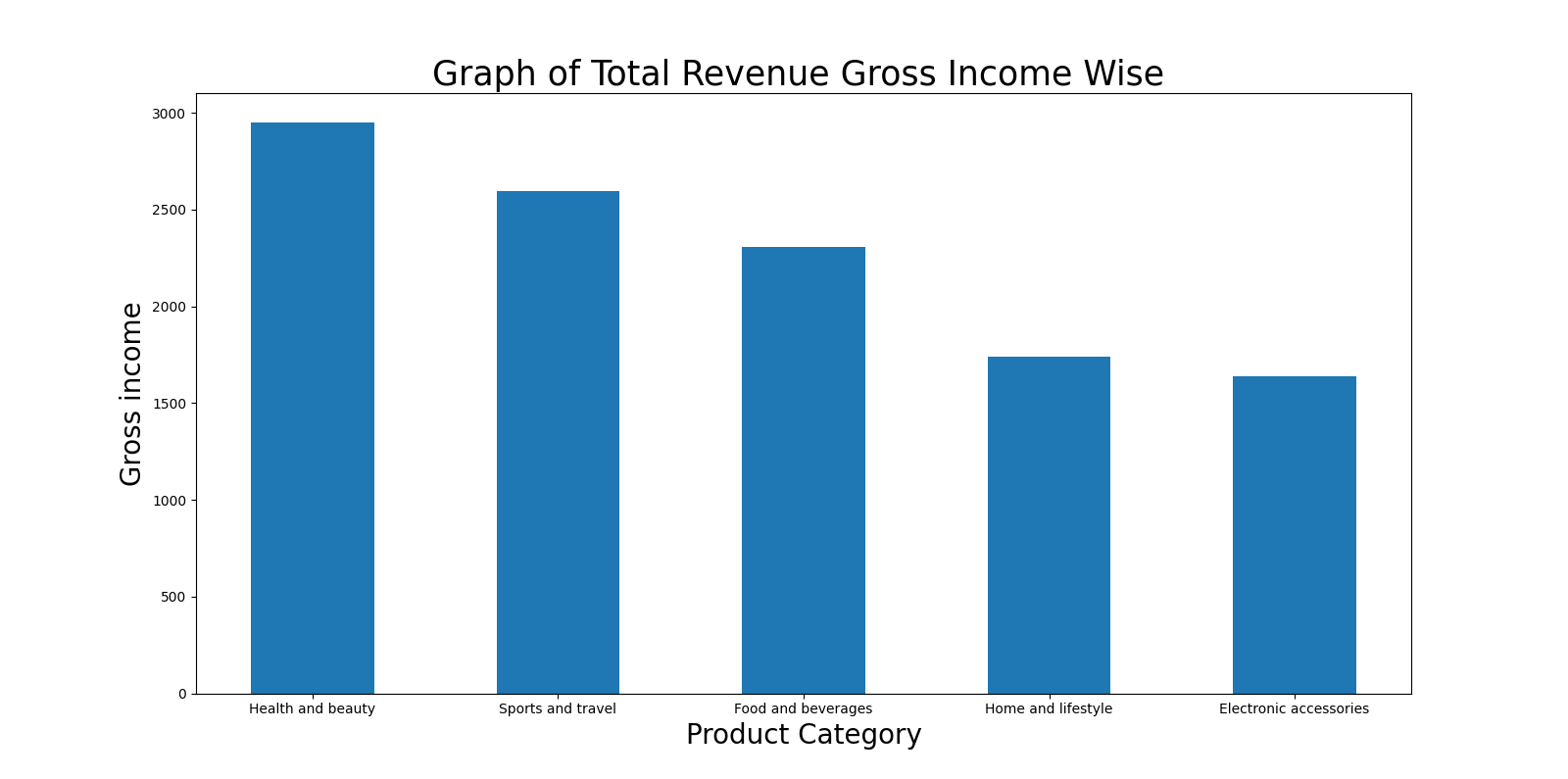
**Figure 9: Output for option 6:(Total Revenue Customer Type Wise)**



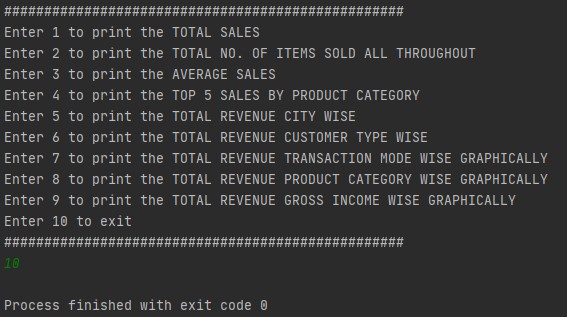
**Figure 10: Output for option 7(Total Revenue Transaction wise)**



**Figure 11: Output for option 8(Total Revenue Product Category Wise)**



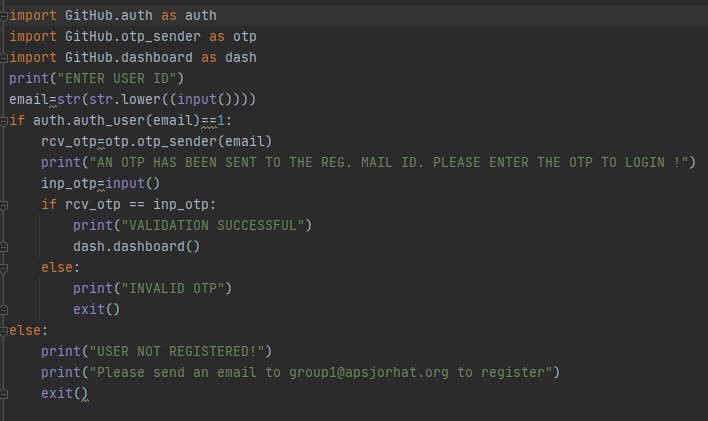
**Figure 12: Output for option 9(Total Revenue Gross Income Wise)**



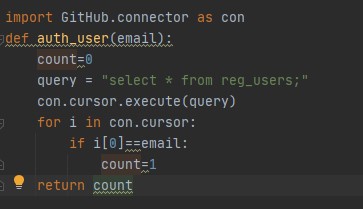
**Figure 13: Output for option 10(Exit)**

**Source Code**

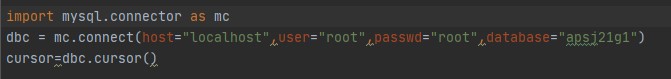
**Main**

****

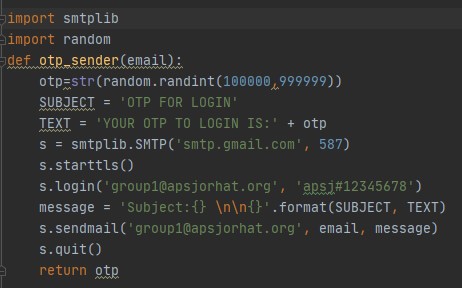
**Authentication**

****

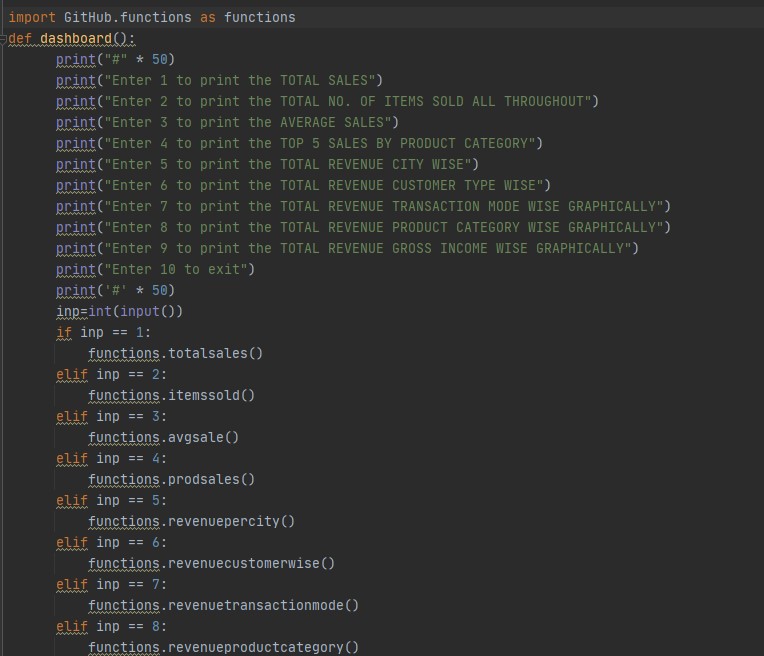
**Connector**

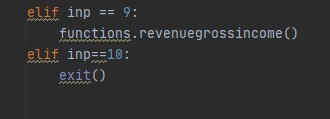
****

**OTP Sender**

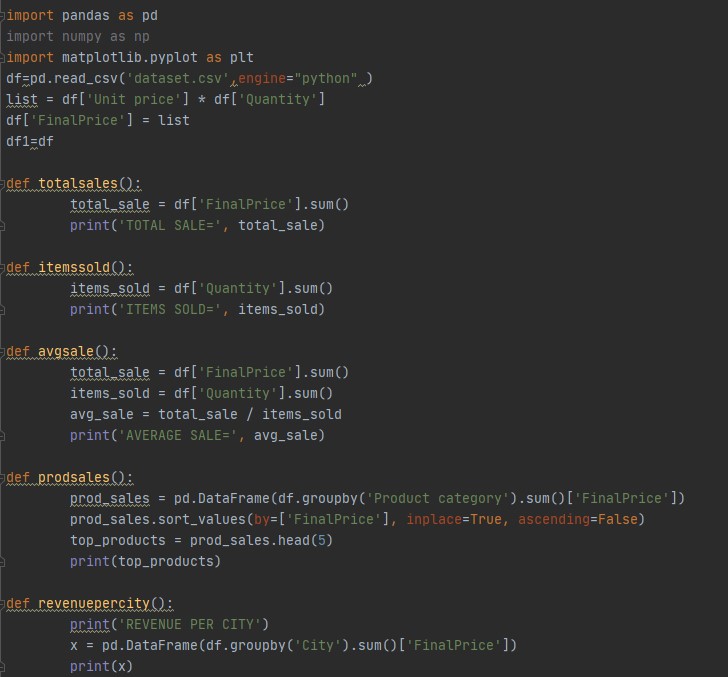
****

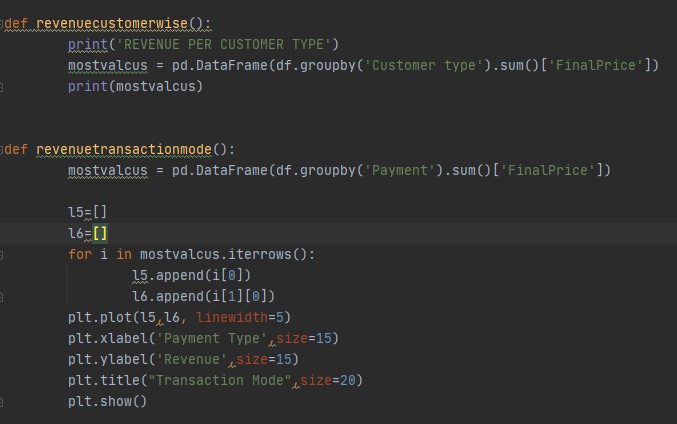
**Dash Board**

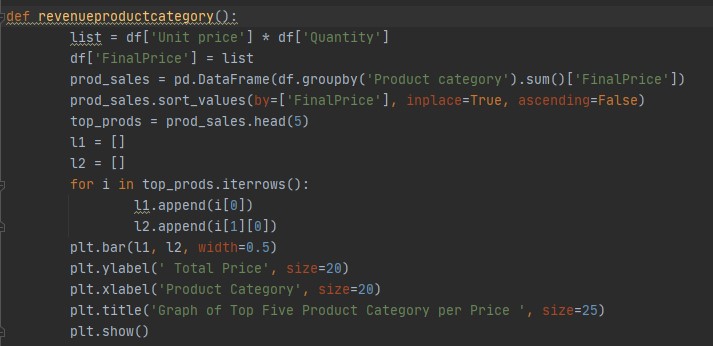
****

****

**Functions**

****

****



**Commands Used in MySQL**

**Creating Database**

Create database apsj21g1;

**Using Database**

Use apsj21g1;

**Creating Table and Inserting Values**

Create table reg\_users (email varchar(30) primary key);

**To Fetch All Values**

Select \* from reg\_users;

**Conclusion and Future Work**

By making this program project, we have successfully shown an example of a useful database, the kind that is used by the companies to extract information. Here we have successfully completed the task of making a database of Amazon inc. And presenting it in front of the user.

The functions performed by this project program are in accordance to our assumptions for further upgradation, we can also add features such as a Graphical User Interface to this project , which can be developed in the form of a web , desktop or mobile application. This project can also be scaled to analyse global data in real time.

**References**

[1] <https://www.limswiki.org/index.php/MySQL>

[2] <https://en.wikipedia.org/wiki/PyCharm>

[3] <https://en.wikipedia.org/wiki/Pandas_(software)>

[4] <https://en.wikipedia.org/wiki/NumPy>

[5] <https://en.wikipedia.org/wiki/Matplotlib>