**SALES AND ANALYSIS**

****

**SUBMITTED BY:**

**KULDEEP MAURYA**

**RANJEEV NAYAK**

**RAJA SHARMA**

**TABLE OF CONTENT**

|  |  |  |
| --- | --- | --- |
| **SNO** | **CONTENT** | **PAGE NO** |
| **1** | **PREFACE** | **2** |
| **2** | **INTRODUCTION** | **3** |
| **3** | **PROJECT DETAILS** | **4** |
| **4** | **ALGORITHM USED WITH EXPLANATION** | **5** |
| **5** | **OUTPUT SCREEN SHOTS** | **6** |
| **6** | **GRAPHS** | **9** |
| **7** | **EXCEL FILE** | **12** |

|  |  |  |
| --- | --- | --- |
| **8** | **SUMMARY** | **16** |

**PREFACE**

The successfully completion of this project was a unique experience

For me because by visiting many Place and interacting various person,

I achieved a better knowledge about sales.

The experience which I gained by doing this project was essential at this turning point of my carrier.

This project is being submitted which content detailed analysis of the research undertaken by me.

The research provide and opportunity to the student to devote his/her skills

knowledge and Competencies required during the technical session.

The research is on the topic **SALES AND ANALYSIS**.

**INTRODUCTION**

* **This project showcases all the steps (from scratch) taken to solve a machine learning problem. For your understanding, I've taken a simple yet challenging data set where you can engineer features at your discretion as well.**
* **We will collect data from stores and make prediction on that gathering data.**
* **From these data we will analyse the profit and loss on that data and analyse that what kind of data will benefit for which season.**

**PROJECT DETAILS**

* **Understand the problem: Before getting the data, we need to understand the problem we are trying to solve.**
* **Data Exploration: It involves creating charts, graphs and cross-tables to understand the behavior of features.**
* **\**Data Preprocessing: \**Here, we impute missing values and clean string variables (remove space, irregular tabs, data time format) and anything that shouldn't be there.**
* **Model Testing: Finally, we test the model on the unseen data (test data) set.**
* **Model Training: Using a suitable algorithm, we train the model on the given data set.**

**Model Evaluation: Once the model is trained, we evaluate the model's performance using a suitable error metric**

**ALGORITHMS USED**

* **Linear Regression**

In statistics, **linear regression** is a **linear** approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X.

* **SVM**

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate

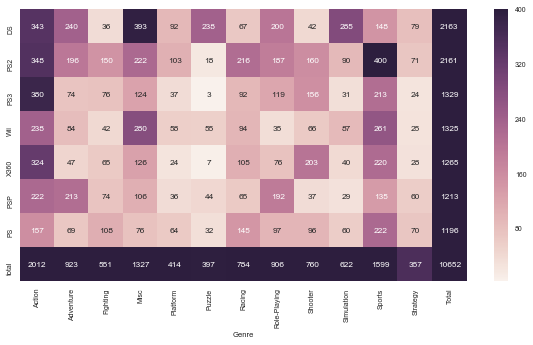
* **KNN**

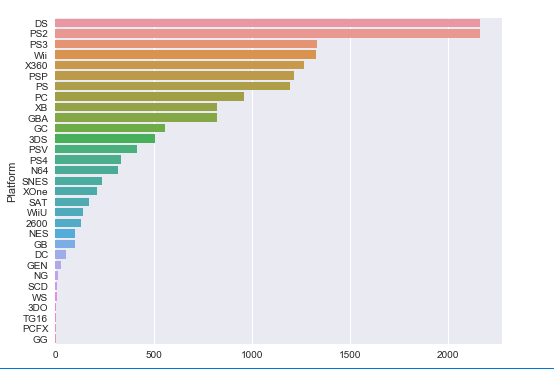
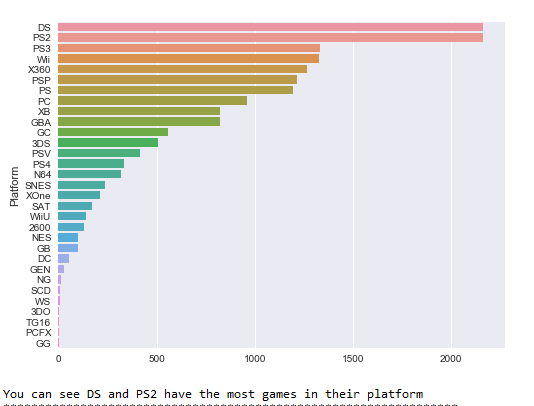
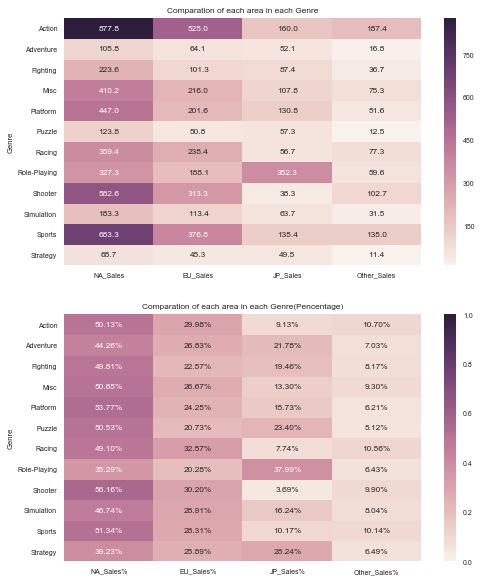
In pattern recognition, the **k**-**nearest neighbors algorithm** (**k**-**NN**) is a non-parametric method used for **classification** and regression. In both cases, the input consists of the **k-**closest training **examples** in the feature space. ... In **k**-**NN classification**, the output is a class membership.

* **K-Means**

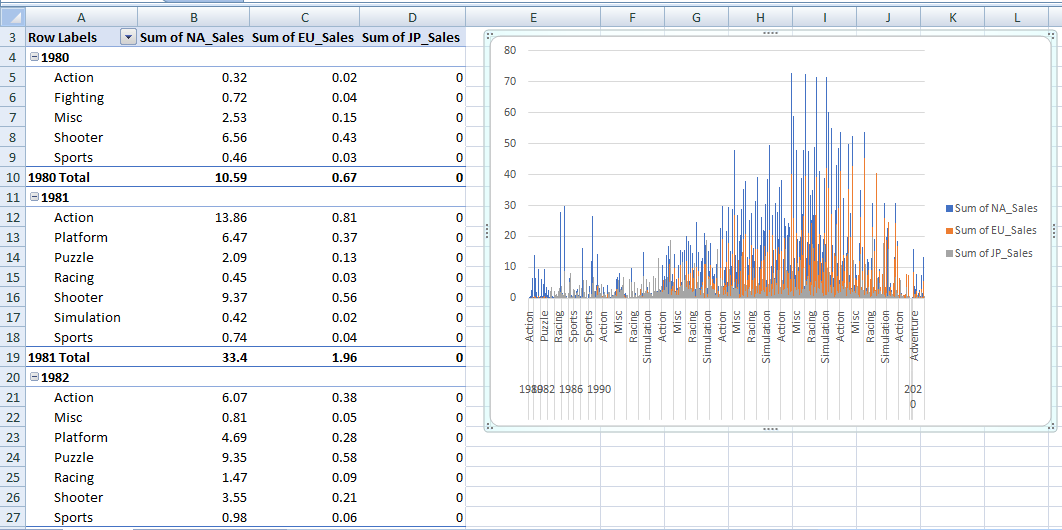
**K**-**means** Clustering in **Python**. **K**-**means** clustering is a clustering algorithm that aims to partition observations into clusters. ... initialization– **K** initial “**means**” (centroids) are generated at random. Assignment – **K** clusters are created by associating each observation with the nearest

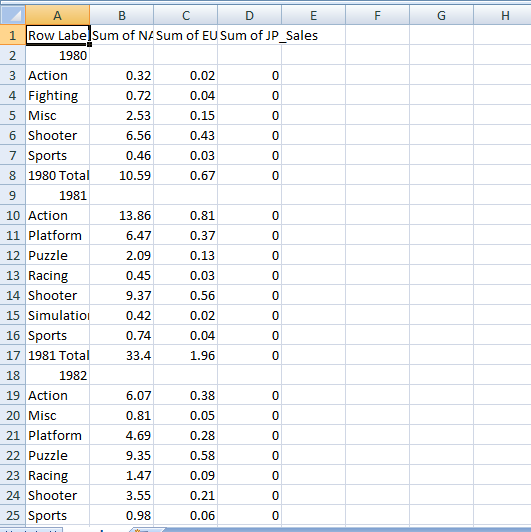
**GRAPHS**

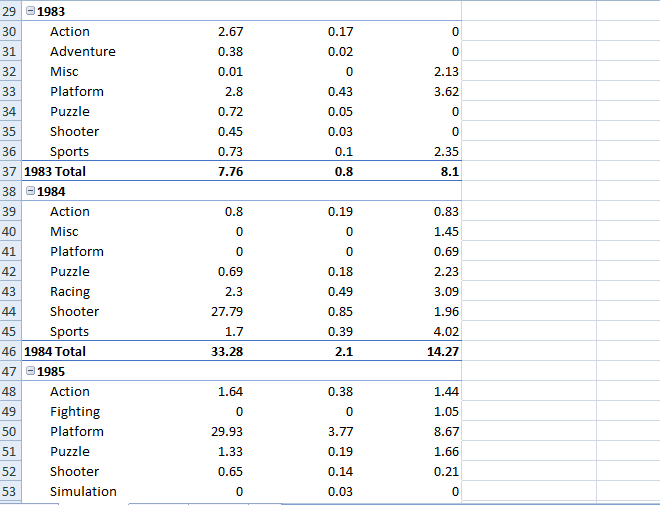
****

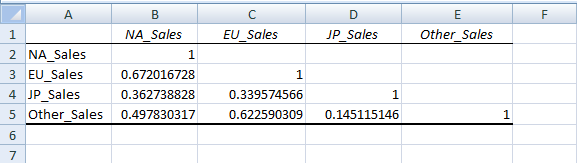
****

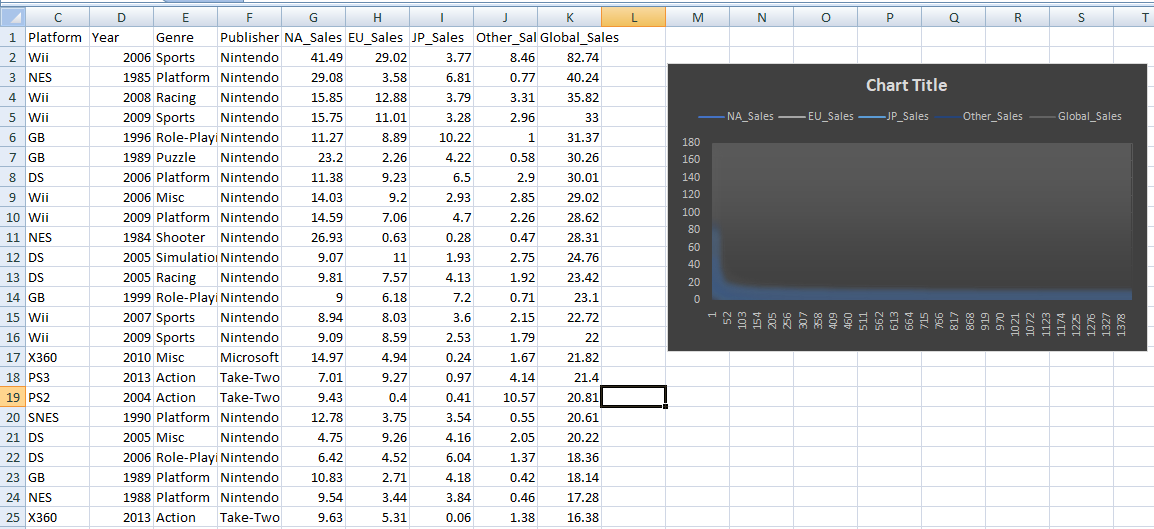
**EXCEL FILE**

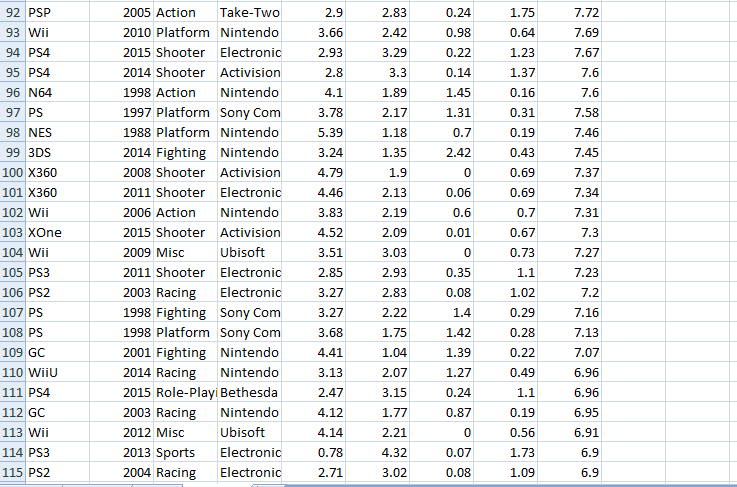
****

****

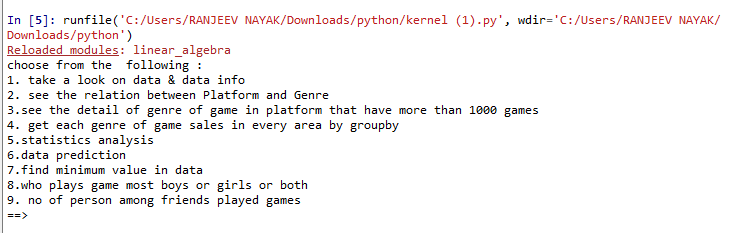
****

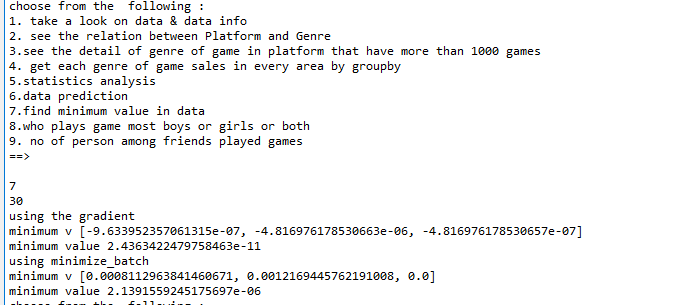
****

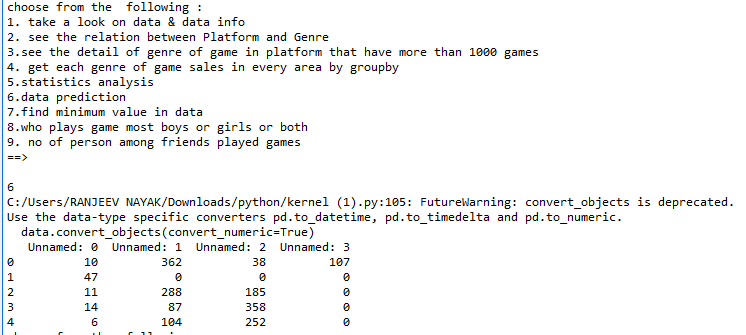
****

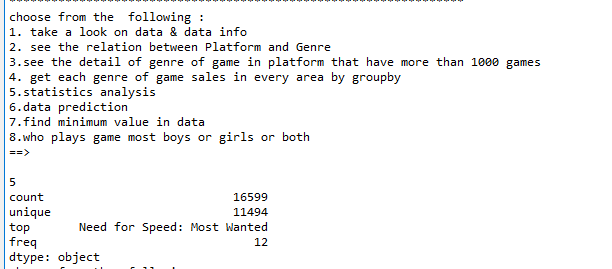
****

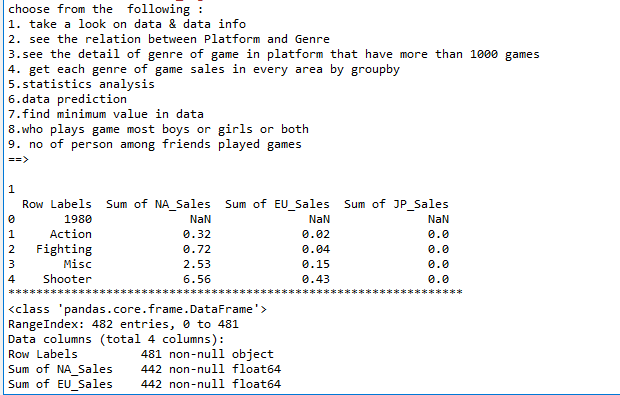
**OUTPUTS**

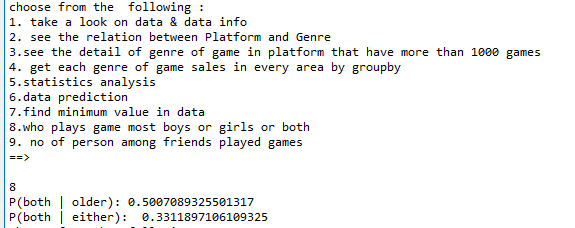
****

****

****

****

****

****

**SUMMARY**

* **We will collect data from stores and make prediction on that gathering data.**
* **From these data we will analyse the profit and loss on that data and analyse that what kind of data will benefit for which season.**
* **Certainly this definitely hit all audience who are interested in playing different**

**Games.**

**This involve kids, teenagers, elders.**

* **Through this project we will find the which genre of games selling most.**