**DATA MANAGEMENT PROJECT REPORT**

(Project Semester: August-December 2021)



**JOB MARKET IN US**

Submitted by

**Sweety Kumari**

**11904066**

Program and Section: B.Tech CSE(hons), KM001

**Course Code: INT217**

Under the Guidance of

**Sandeep Kaur :23614**

Discipline of CSE/IT

**Lovely School of Computer Science & Engineering**

**Lovely Professional University, Phagwara**

**CERTIFICATE**

This is to certify that **SWEETY KUMARI** bearing Registration no. **11904066** has completed **INT217** project titled, “**JOB MARKET IN US**” under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort and study.

**Sandeep Kaur**

**School of Computer Science & Engineering**

**Lovely Professional University**

**Phagwara, Punjab.**

**DECLARATION**

I, Sweety Kumari, student of B.Tech CSE under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

**Registration no.:-11904066 Sweety Kumari**

**ACKNOWLEDGEMENT**

A project work is a combination of views, ideas, suggestions and contribution of many people. Thus, one of the pleasant parts of writing the report is to thank those who have contributed towards its fulfilment.

I consider it as great privilege to have esteemed Lecturer Ms. SANDEEP KAUR as my project guide. I take this opportunity to express my sincere gratitude to her through constant advice and constructive criticism nourished my interest in the subject and provided a free and pleasant atmosphere to work against all odd situations. I avail this opportunity to extend my heart full thanks and deep respect to faculty member for their able guidance during this project.

My gratitude to all those, who **responded to my questionnaire** in a well-defined manner and helped me acquiring knowledge.

I would like to communicate a deep sense of gratitude to all these people without whom my project would not have been such a great learning experience.

**Sweety Kumari**

**Reg no- 11904066**

**TABLE OF CONTENT**

**Sr.no page no.**

1. Introduction 6

2. Scope of the Analysis 7

3. Source of dataset 7

4. ETL process 8

5. Analysis on dataset (for each analysis) 10

i. Introduction

ii. General Description

iii. Specific Requirements, functions and formulas

iv. Analysis results

v. Visualization

6. List of Analysis with results 16

7. References 21

8. Bibliography 21

**INTRODUCTION**

**Data analysis** is the process of collecting, modelling, and analysing data to extract insights that support decision-making. There are several methods and techniques to perform analysis depending on the industry and the aim of the analysis.

All these various methods for data analysis are largely based on two core areas: quantitative methods and qualitative methods in research.

Gaining a better understanding of different techniques for data analysis, and methods in quantitative research as well as qualitative insights will give your information analyzing efforts a more clearly defined direction, so it’s worth taking the time to allow this particular knowledge to sink in.

DATA ANALYSIS PROCESS:

Data Analysis consists of the following phases

* Data requirements gathering
* Data collection
* Data cleaning
* Data analysis
* Data interpretation
* Data visualization

**INTRODUCTION TO DATASET**

* We will compare every company and there opening for and total number of people applied for the jobs
* We will check most of the people applied for which position and low position also check why people want those jobs
* We will analyze and put a solution why most of people are moving towards that field rather leaving the old traditional field

**SCOPE OF ANALYSIS**

This project on Job market Sheet which provides the overall Statistics details of the different- different job in the market .

**Objectives of this project:**

* To good hand on excel.
* To use different feature and get friendly with the excel.
* To learn the ETL process in the Tableau Prep.
* How to link one sheet to another and traverse between different sheets.
* How to use pivot table and pivot chart.
* Learn to make dashboard in excel.
* To make different type of graphs in excel.
* To learn how to fetch data from other source to excel in different format

Aim of this project is to answer the above objectives in the form of visualization by creating a dashboard to convey the answers effectively and efficiently.

**SOURCE OF DATASET**

The dataset is taken from the Kaggle with the name ‘JOB MARKET IN US’.

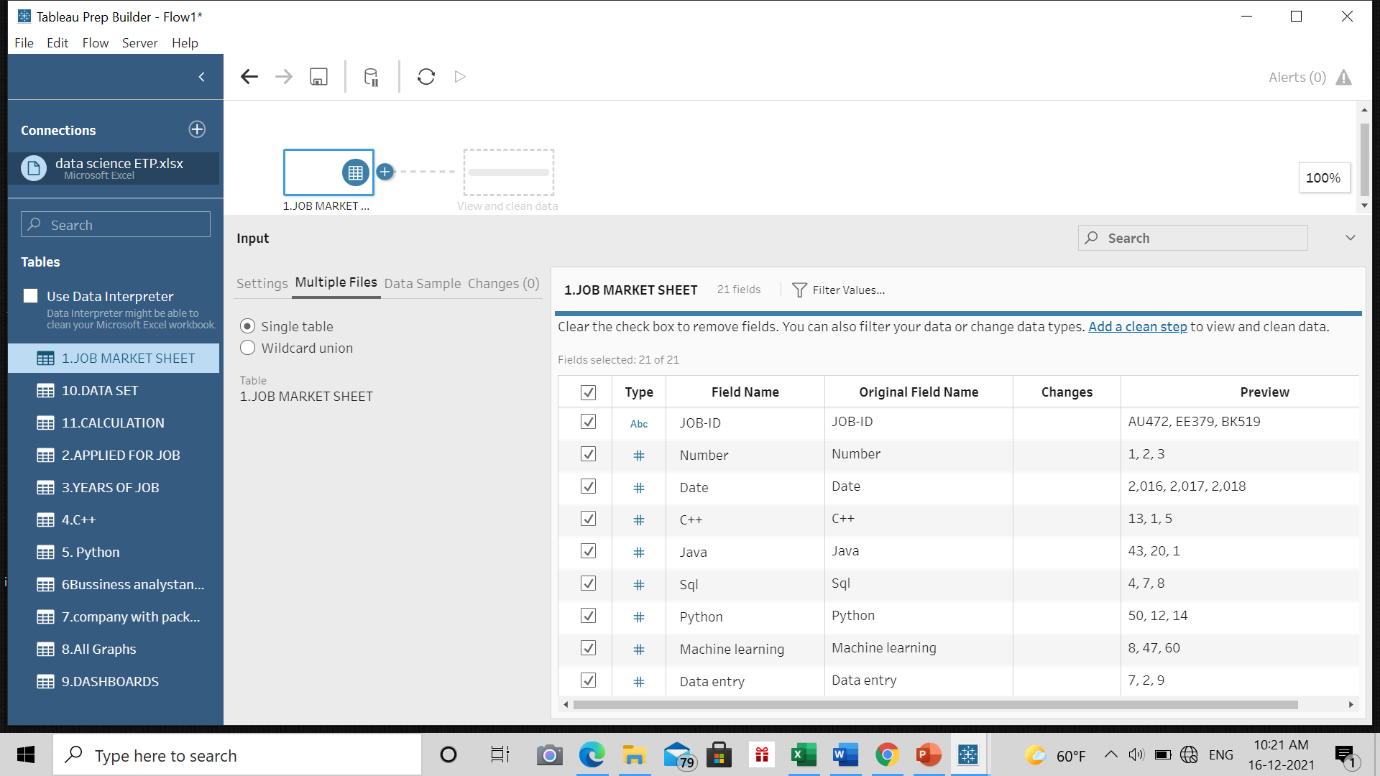
* [**https://www.kaggle.com/sl6149/data-scientist-job-market-in-the-us#fulltimeSEA.csv**](https://www.kaggle.com/sl6149/data-scientist-job-market-in-the-us)
* [**https://www.kaggle.com/datasets?search=data+science+jobs**](https://www.kaggle.com/datasets?search=data+science+jobs)

**ETL PROCESS**

In computing, extract, transform, load (ETL) is a process in database usage to prepare data for analysis, especially in data warehousing. Data extraction involves extracting data from homogeneous or heterogeneous sources, while data transformation processes data by transforming them into a proper storage format/structure for the purposes of querying and analysis; finally, data loading describes the insertion of data into the final target database such as an operational data store, a data mart, or a data warehouse. A properly designed ETL system extracts data from the source systems, enforces data quality and consistency standards, conforms data so that separate sources can be used together, and finally delivers data in a presentation-ready format so that application developers can build applications and end users can make decisions.

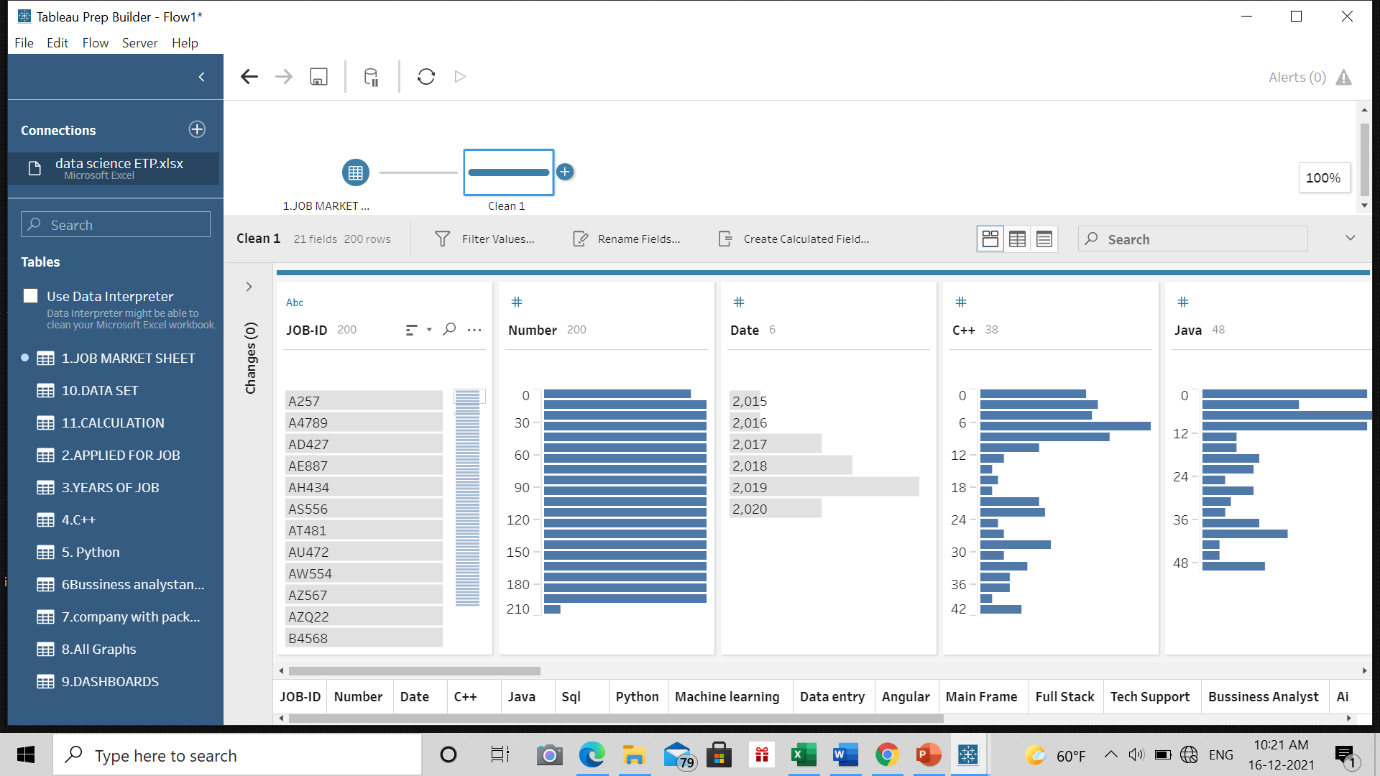
Precisely, ETL is defined as a process that extracts the data from different RDBMS source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system. ETL stands for Extract, Transform and Load. Before ETL.

This data is taken from **Kaggle.**

****

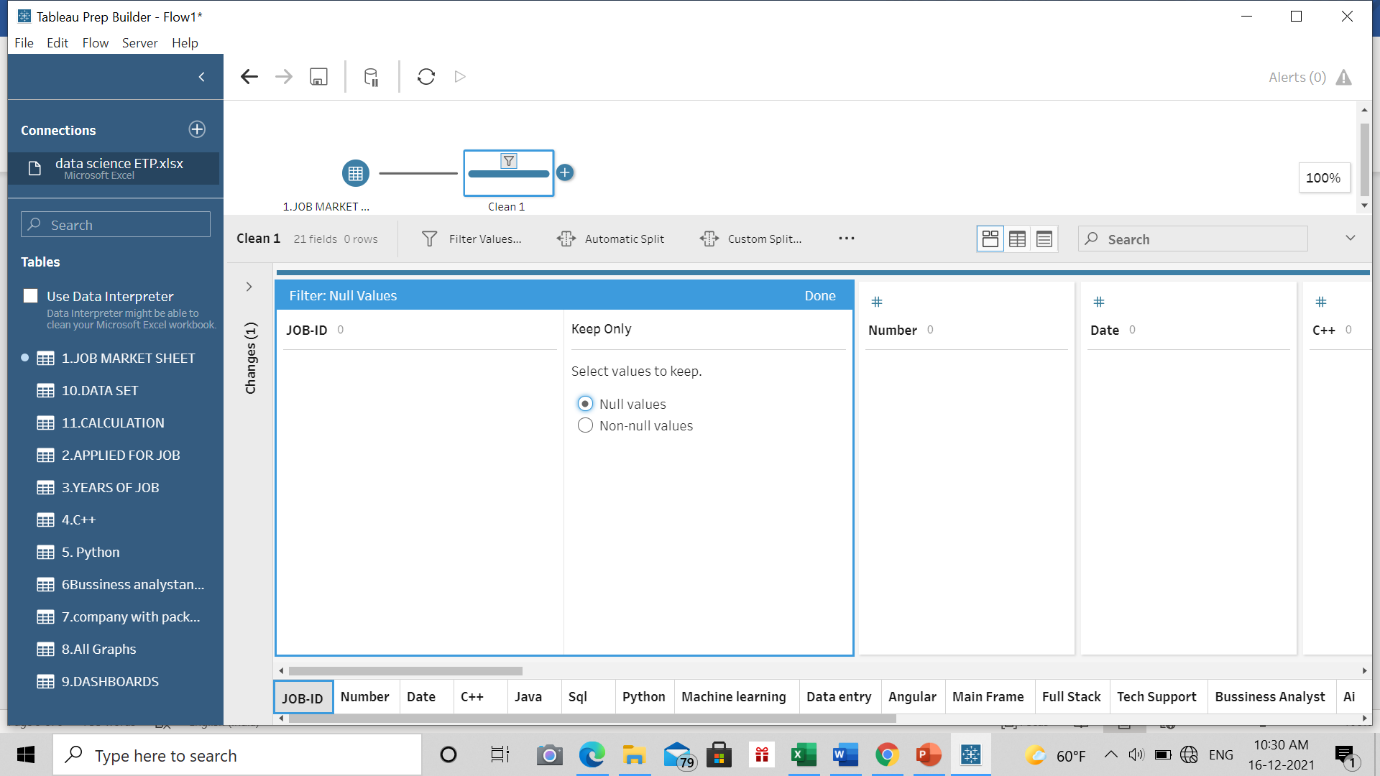
Through the process of ETL, we are going to cleanthe dataset.

**Cleaning the data set:-**

****

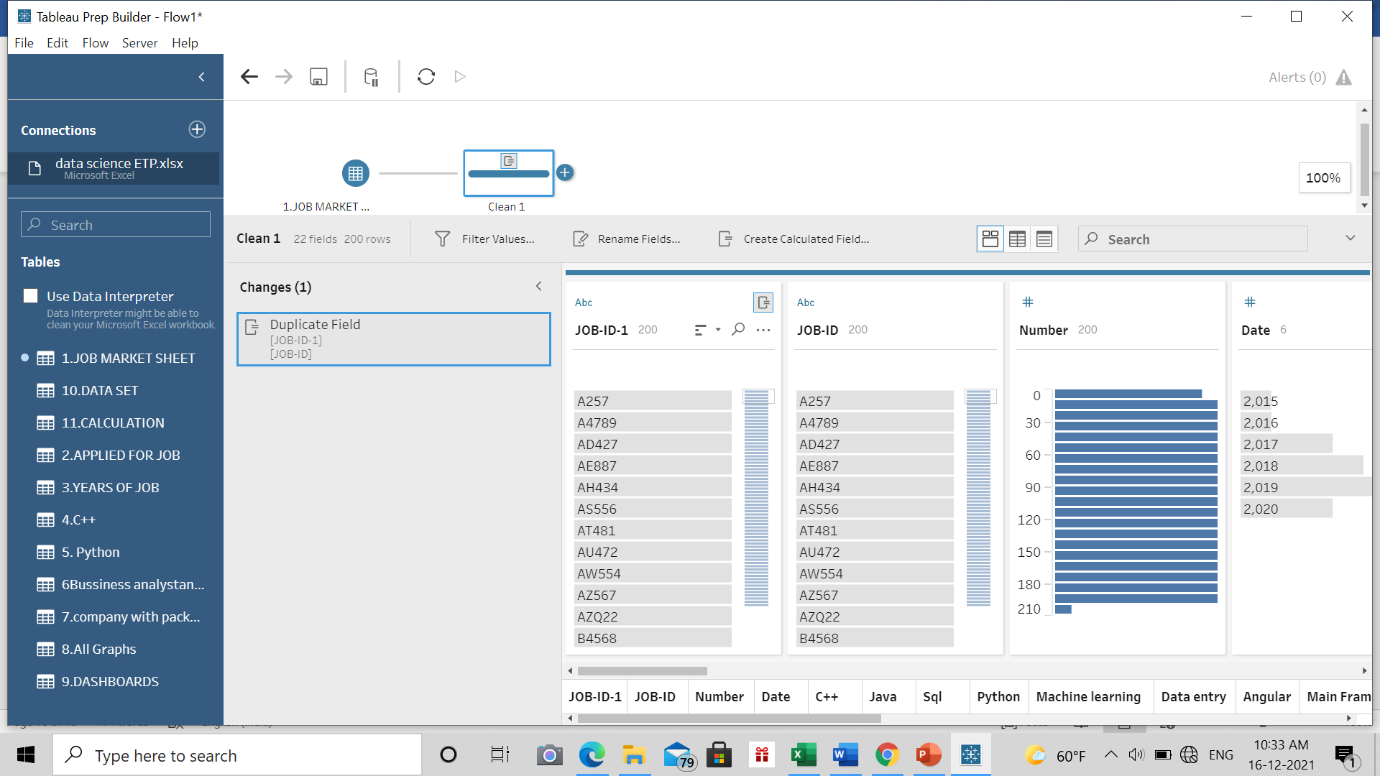
**Excluding the NULL values from the data:-**

We’ll be using Tableau prep for this work as it’ll make the work simple and faster because we might not know how many null values could be there in this huge data set. Tableau helps us doing one step cleaning with ease.

****

**Removing Duplicate Values:-**

It might be possible that our data may be containing duplicate values which may hinder in precise analysis. So, our last task in ETL will be removing duplicate values and making our data perfect for analysis.

****

**ANALYSIS OF DATASET**

1. **Total number of people applied for jobs.**

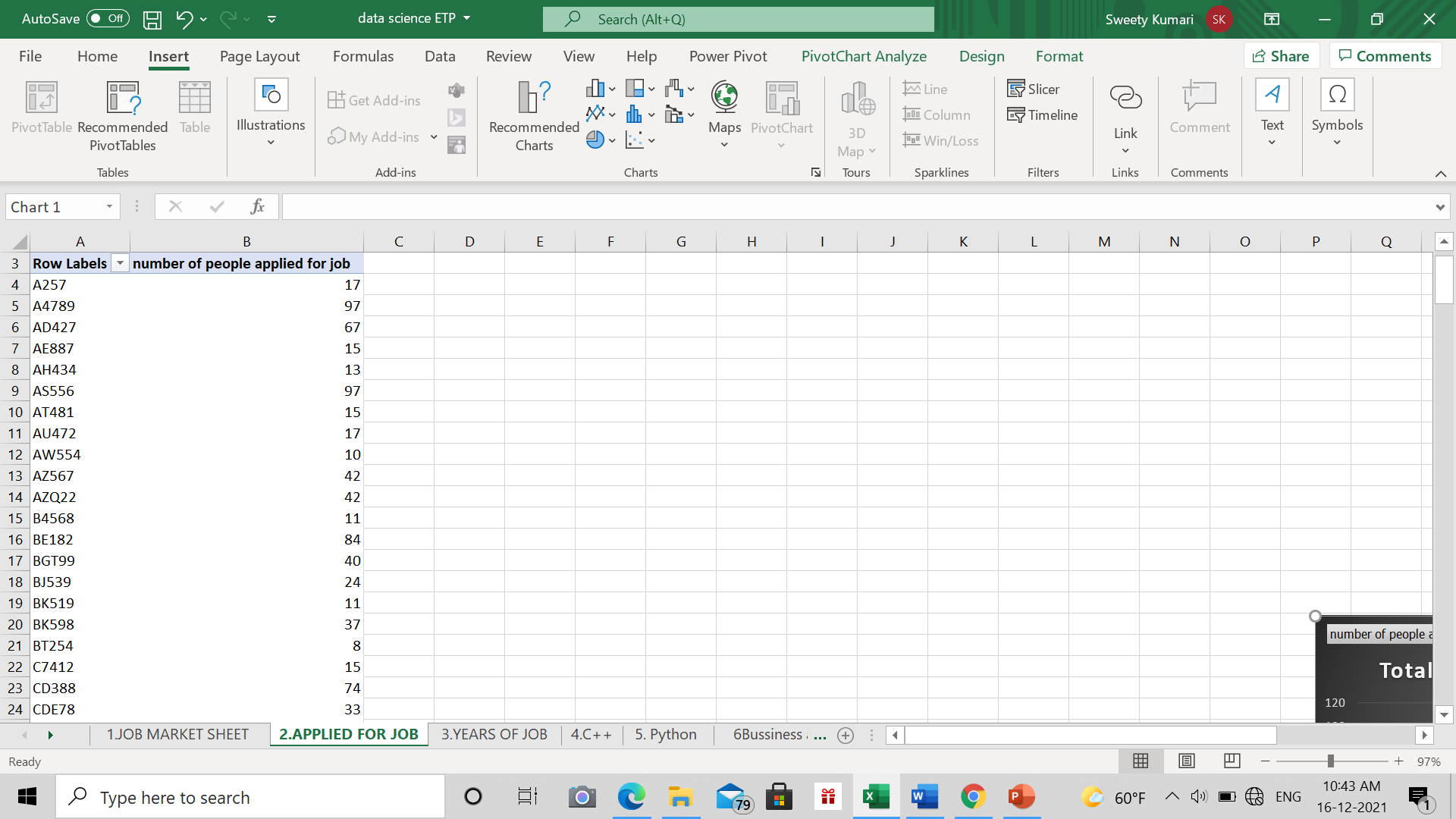
**Description**

In this objective we will be finding people applied for the jobs for the various companies with different role in the companies with their job ids.

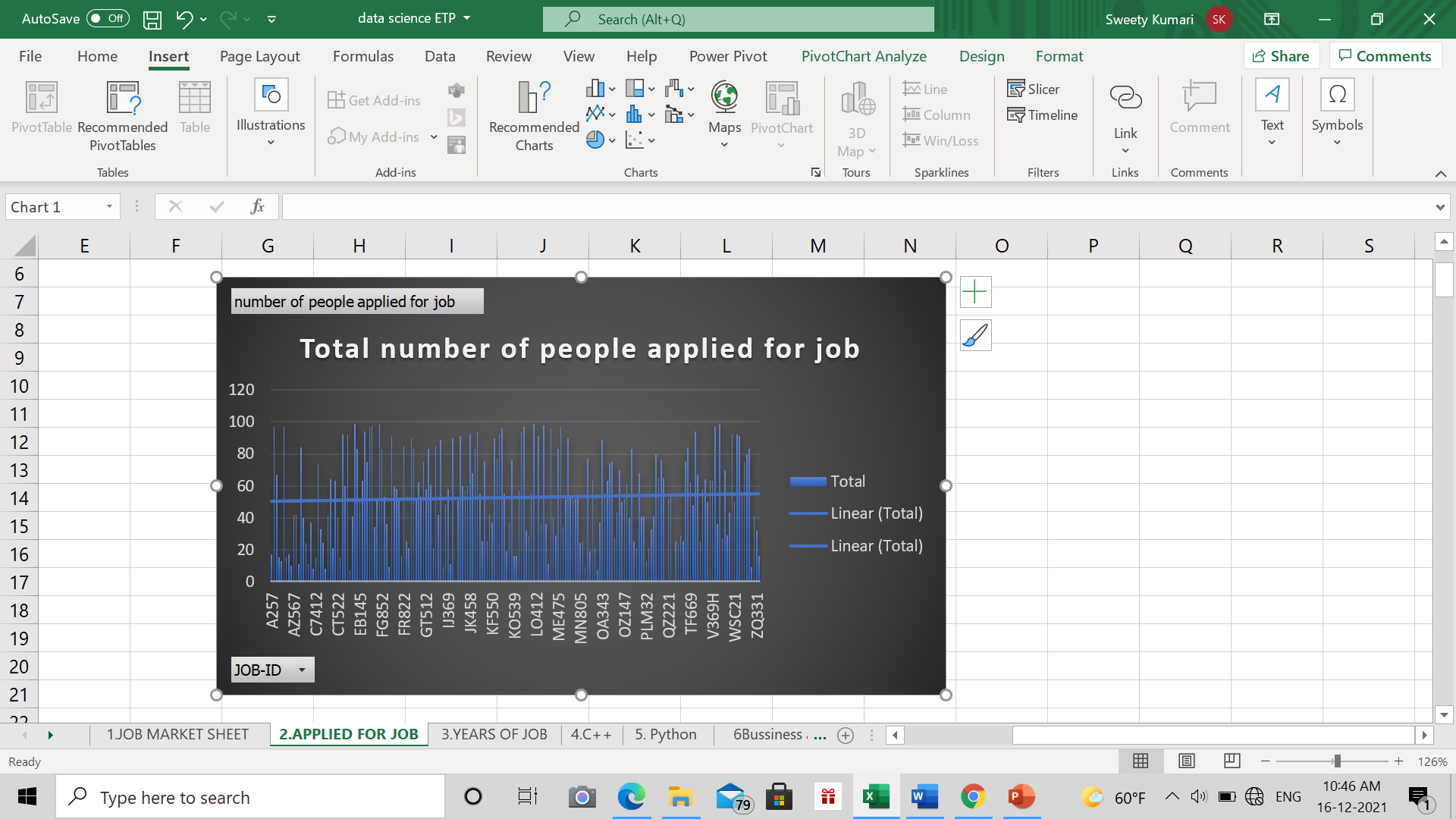
**Specific function and requirements**

We have to create a pivot table to determine the number of people who applied for job and then visualize it on graph.

**Results**

****

**Visualization**

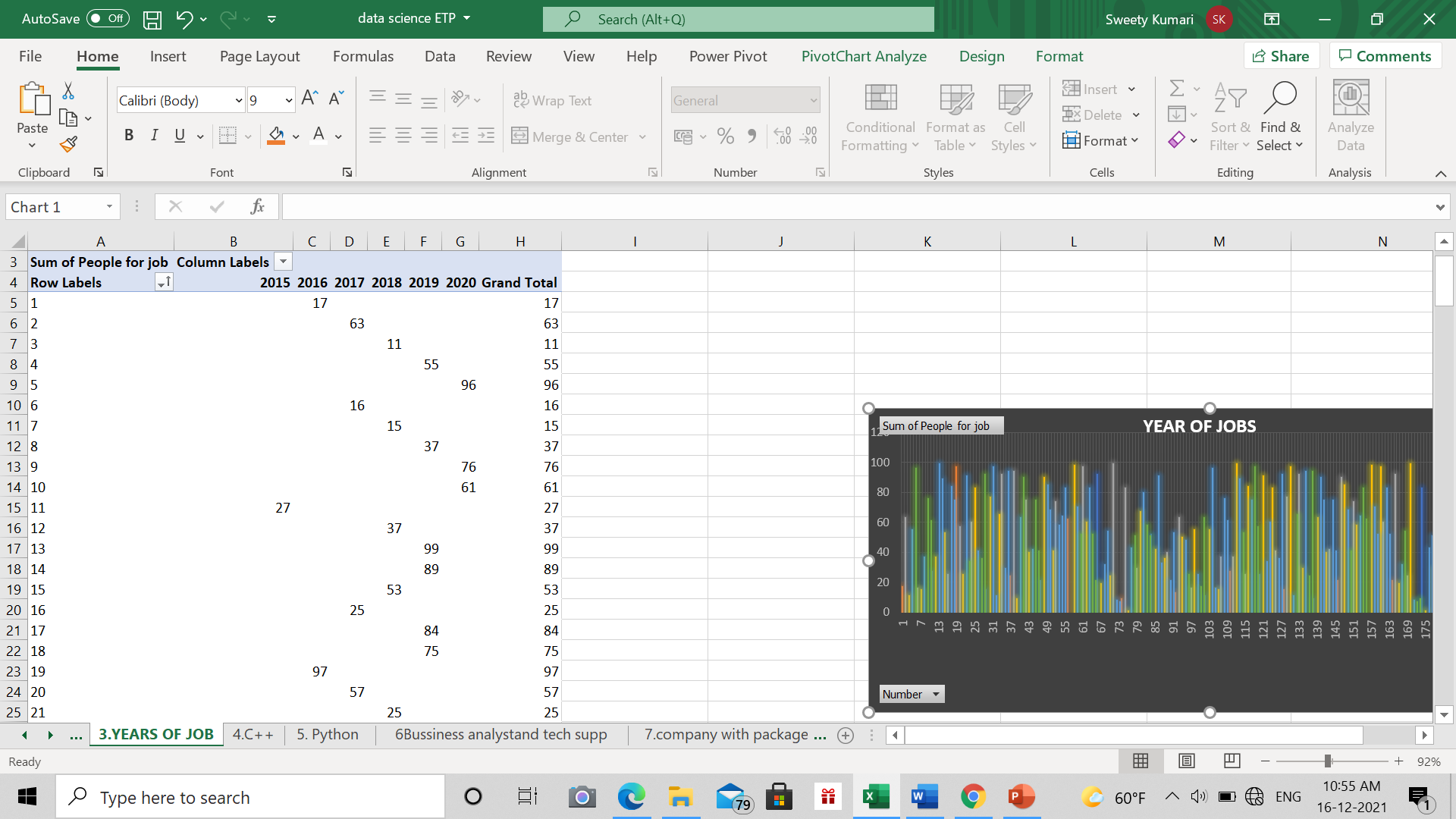
****

1. **Year of job**

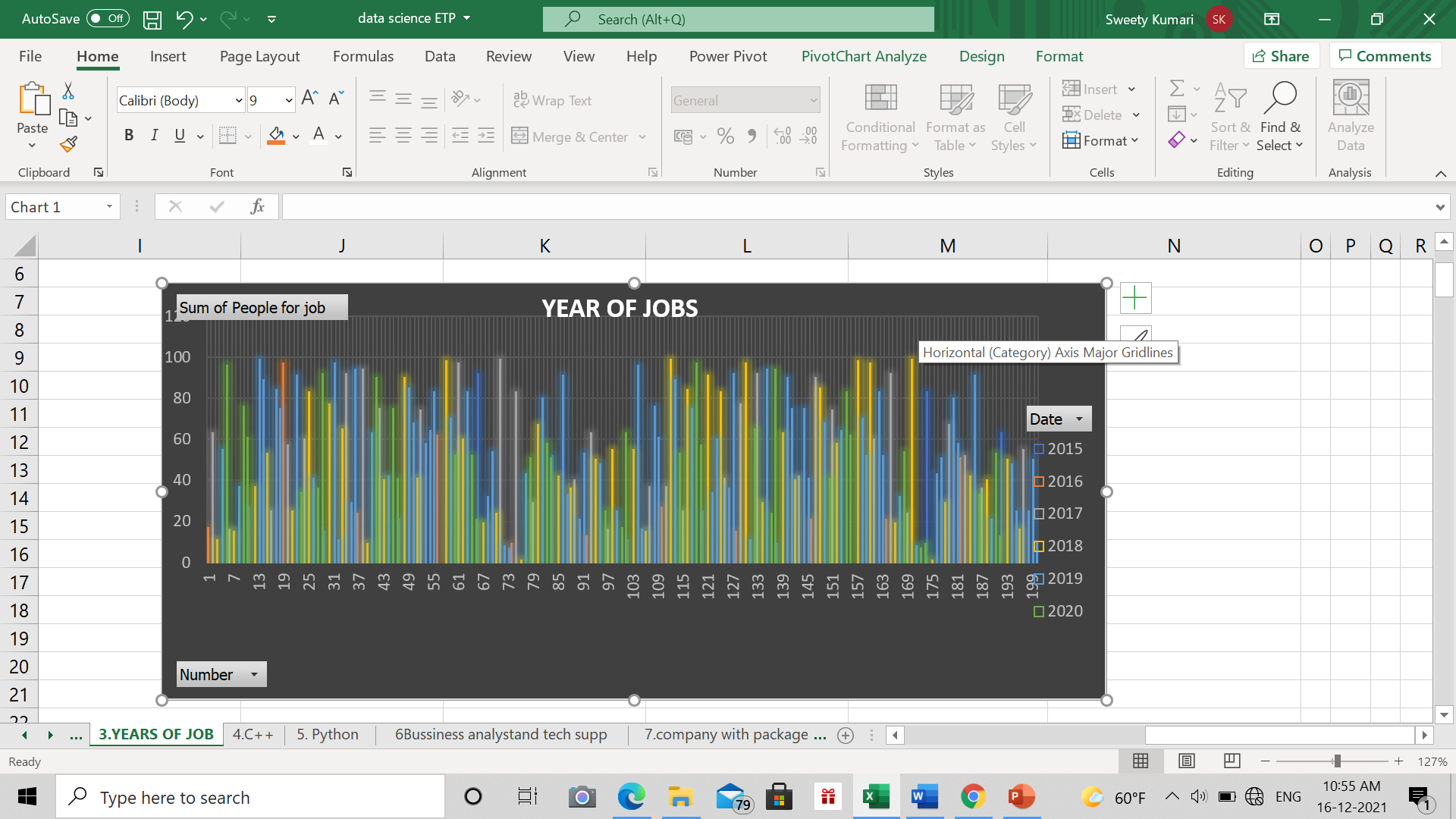
**Description**

In this objectiveshows the year of hire done on every year on various companies from 2015 to 2020 it has the hire details of every year by looking at the graph we can see the hire done for each year

**Results**

****

**Visualization**

****

1. **Applied for C++**

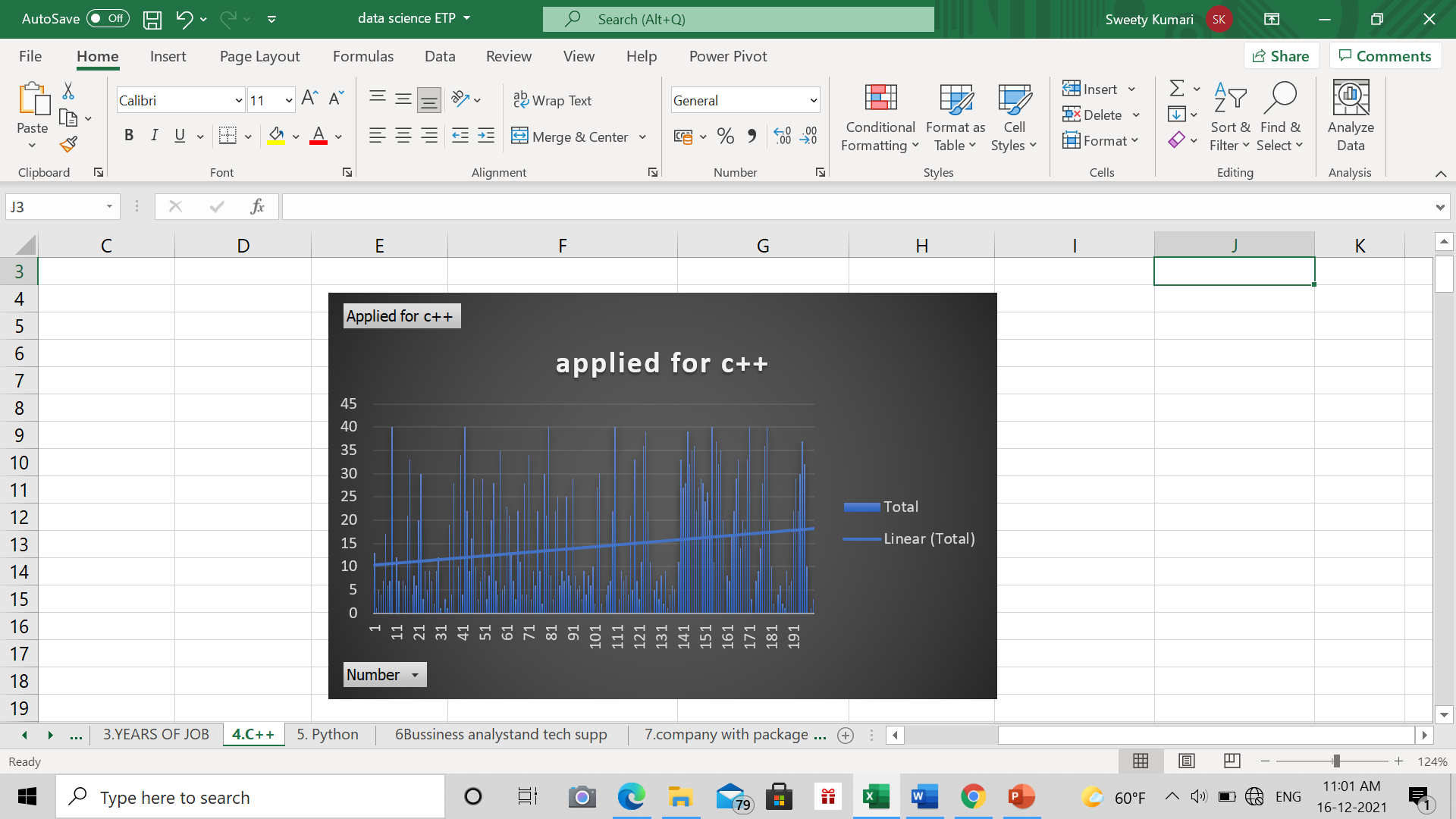
**Description**

In this objective the job applied for c++ role we see very less number of people had applied for this role in various companies.

**Results**

****

**Visualization**

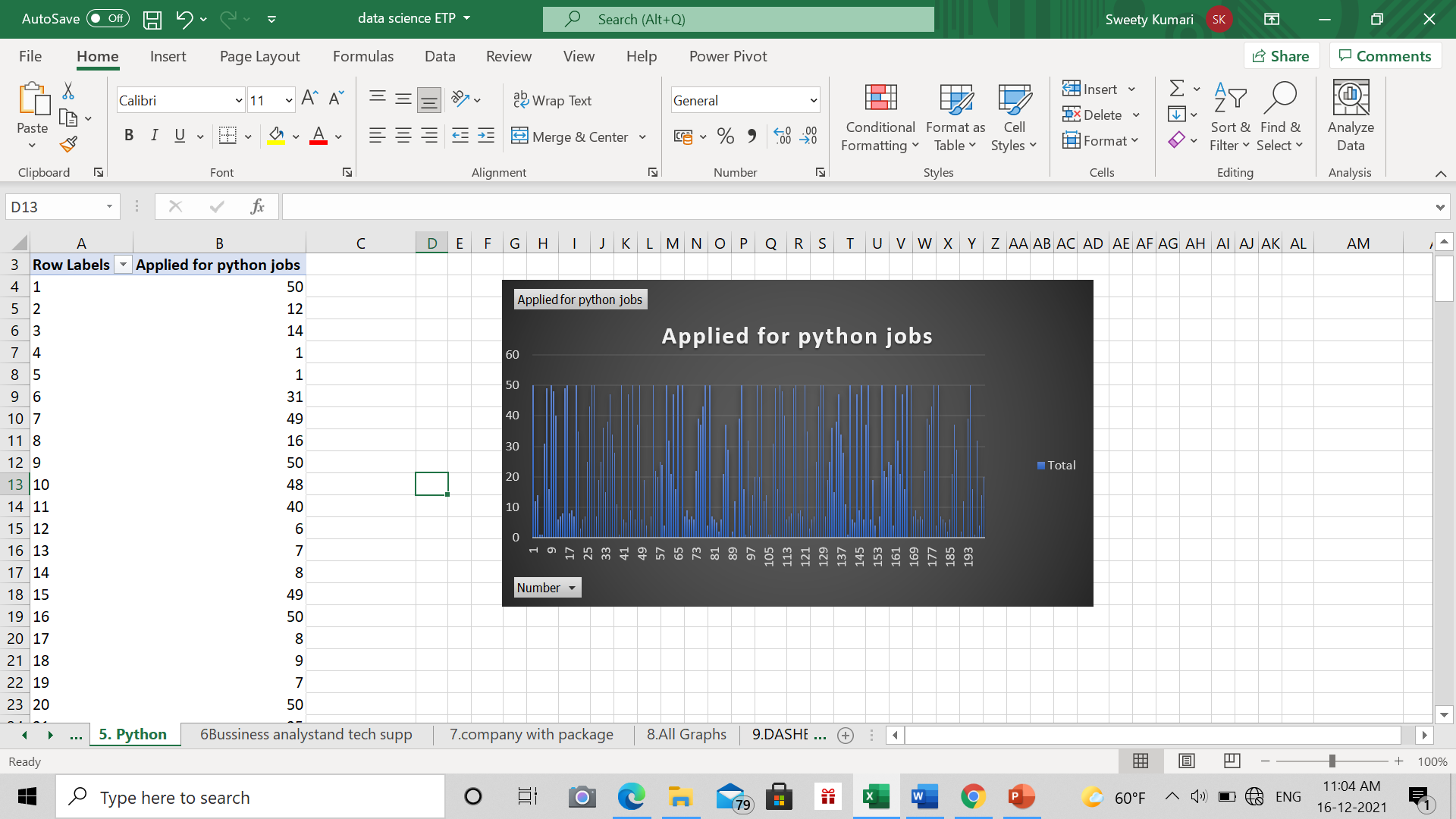
****

1. **Applied for python**

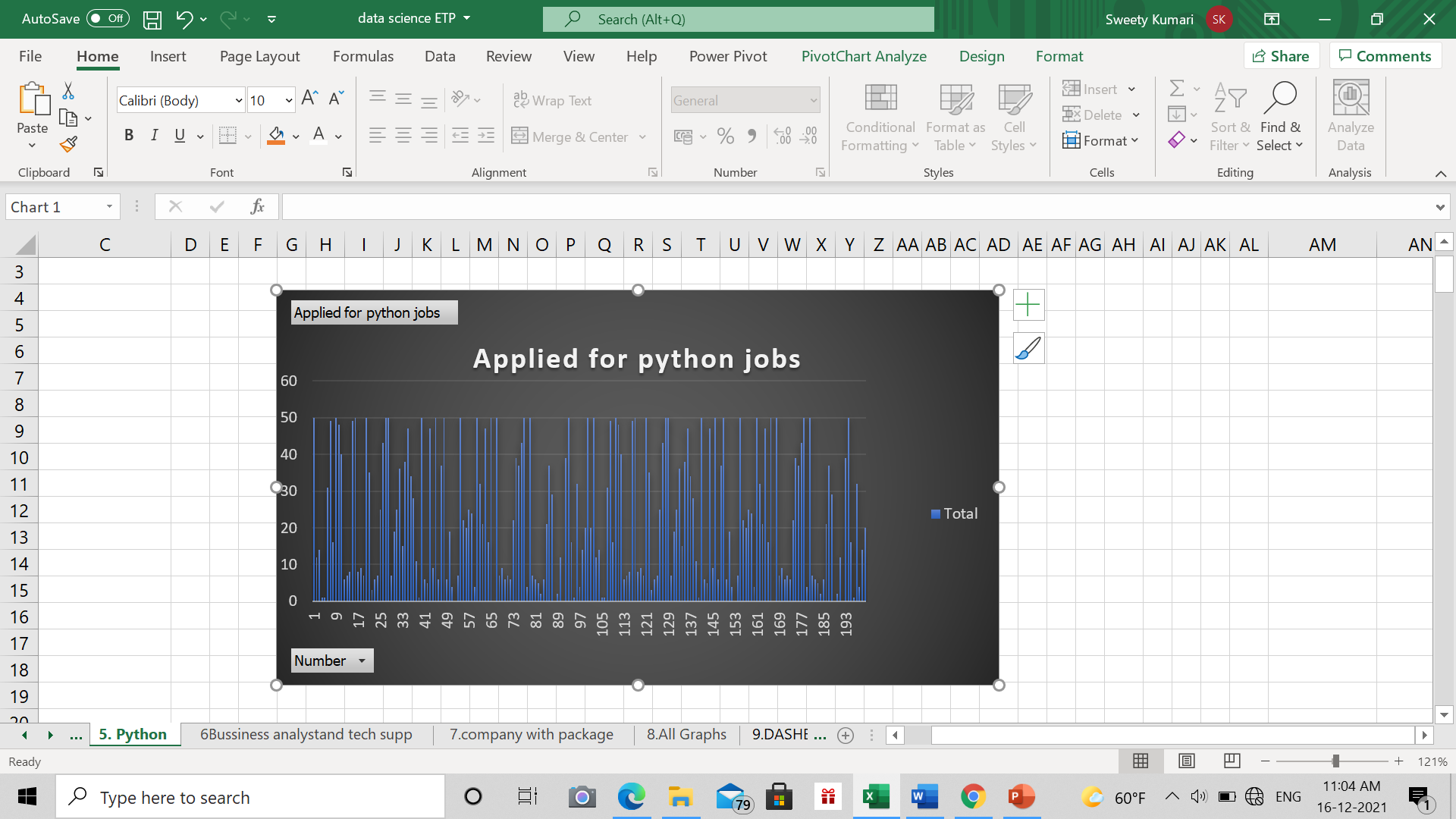
**Description**

In this objective The python is most powerful growing programming language where most of tech companies uses it lot like   
Facebook , Netflix.  
  
from the graph we have seen that many applicants on the python developer role for various companies

**Results**

****

**Visualization**

****

1. **Applied for business analytic and tech support**

**Description**

In this objective Business analyst job used to focus on company business shares…   
the Ba job is also most growing job in the job market companies like delloite,broad reach. Many financial tech companies want this ba roles

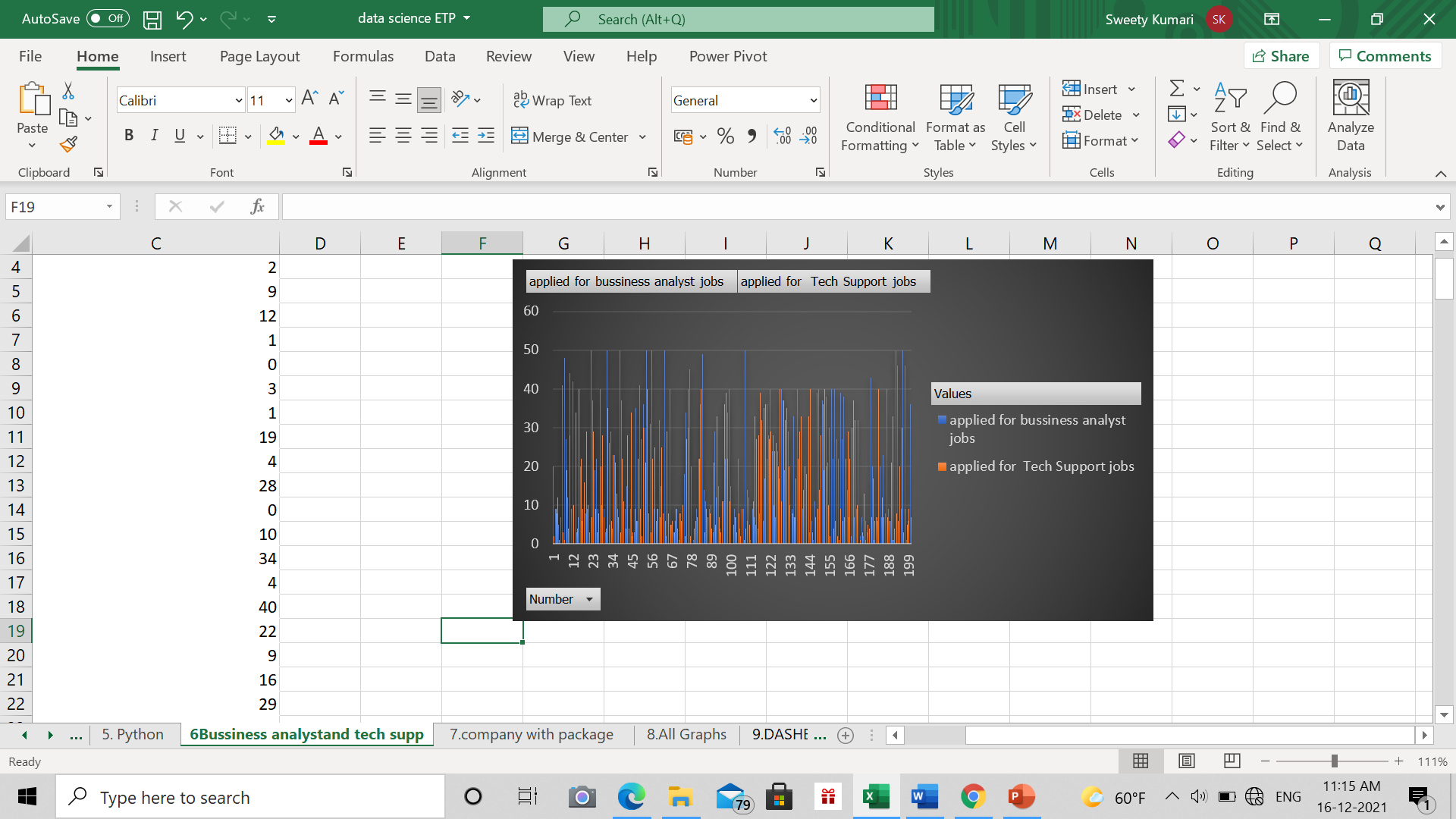
The tech support is technical solving or write a complain type of job used in various aspects  
companies like amazon flip kart has these major demand role

From the graph we have seen the job applicants of ba roles and tech support.

**Results**



**Visualization**

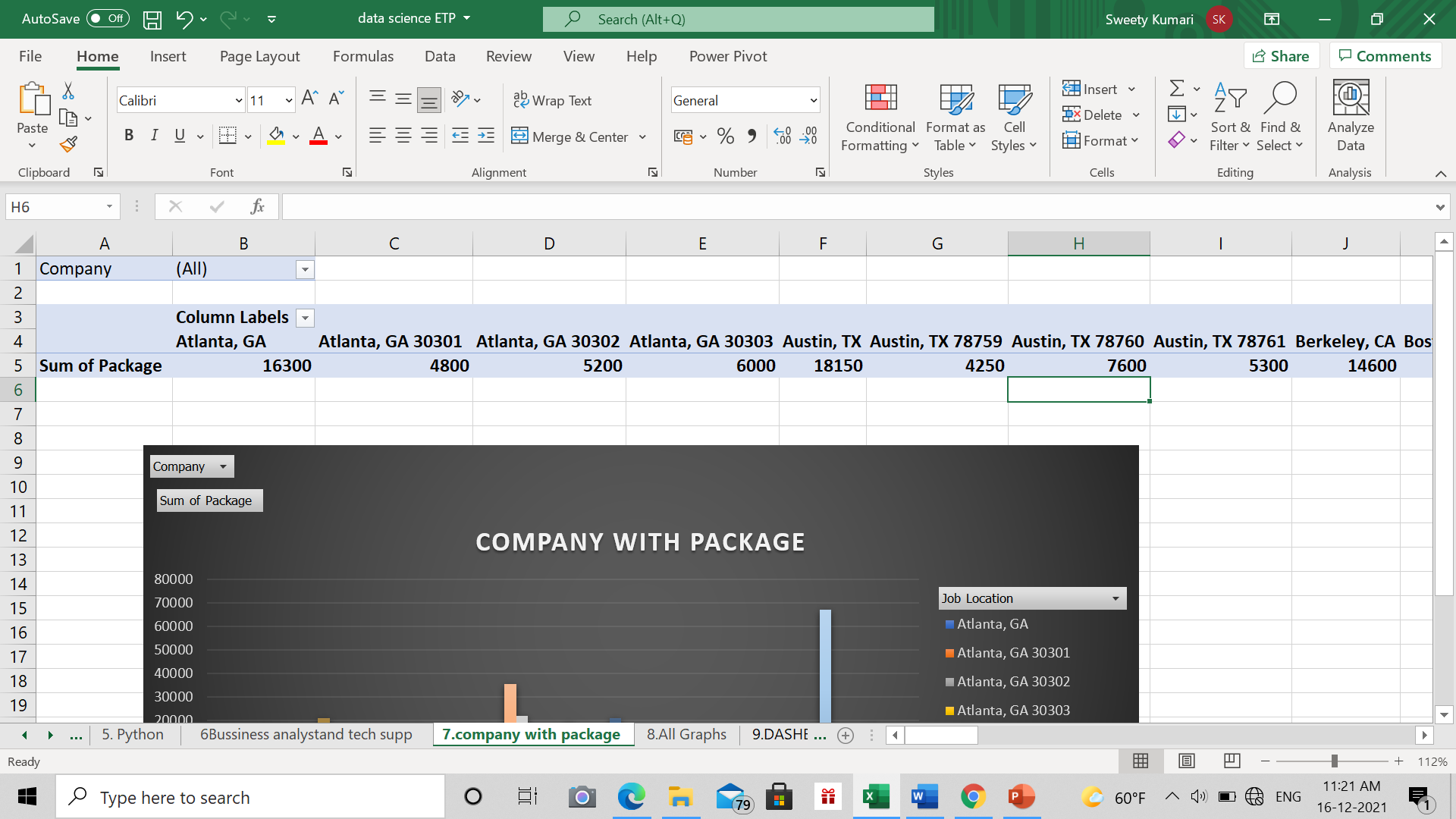


**6.Company with package**

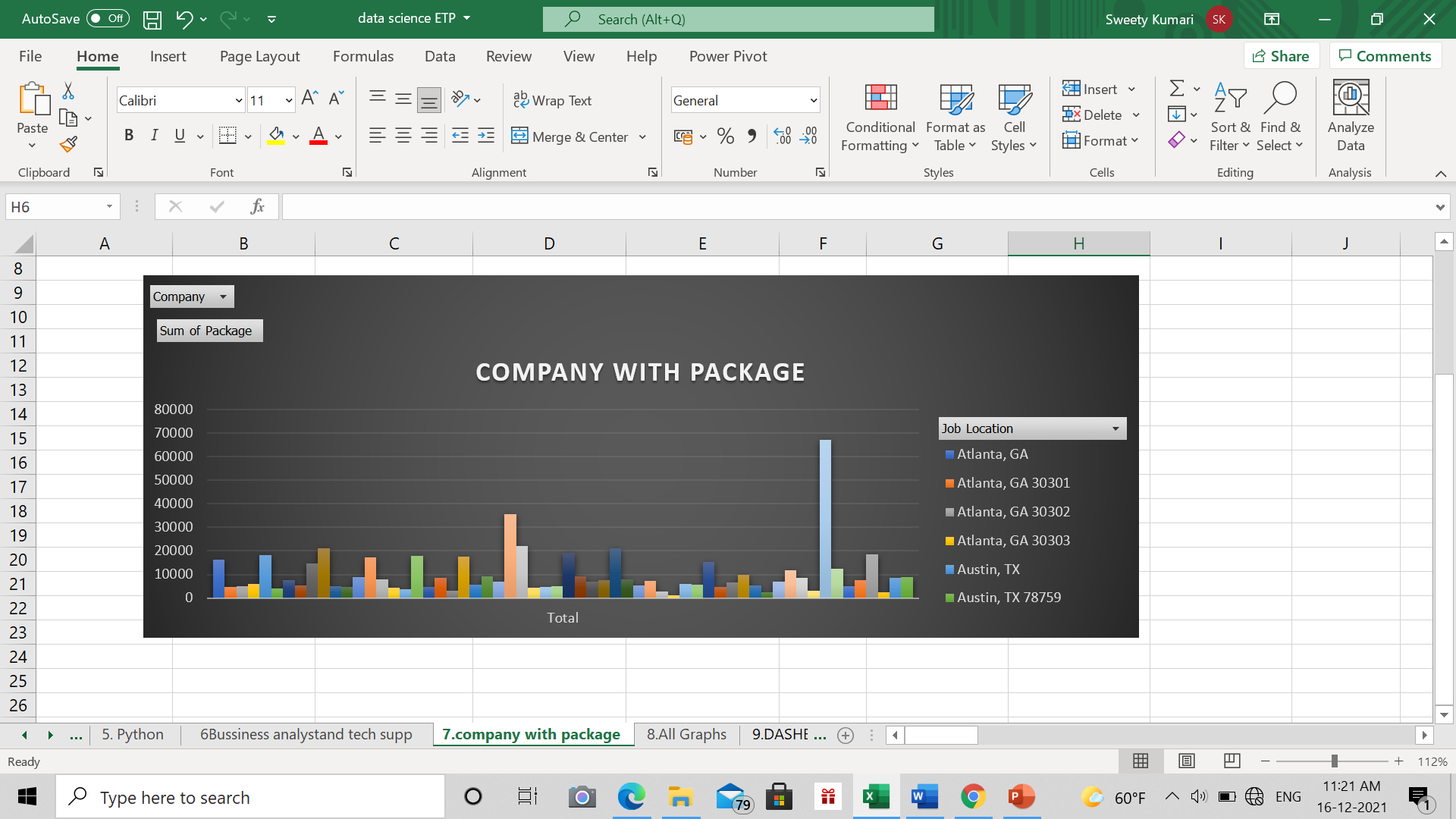
**Description**

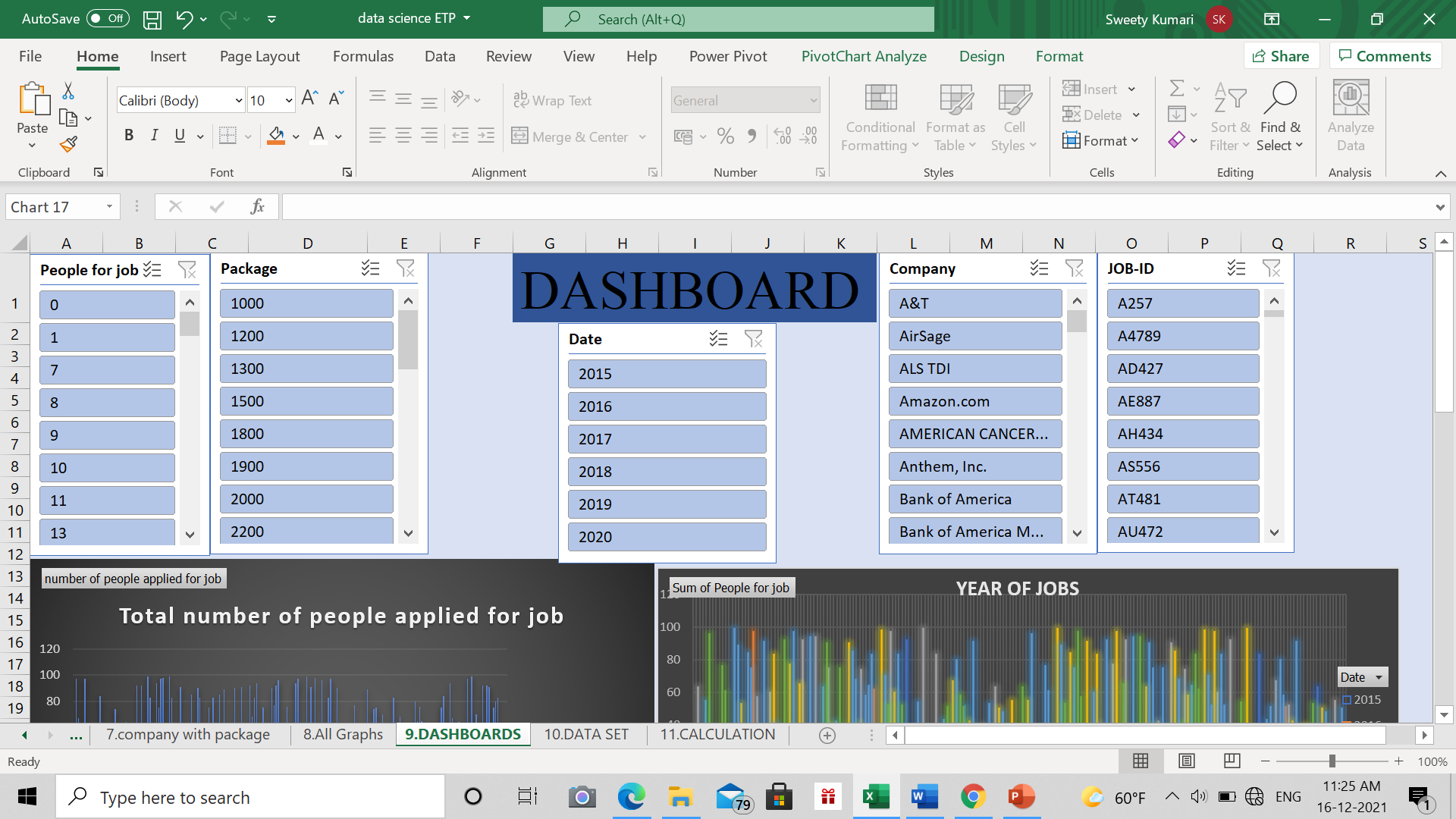
In this objective we will be finding which company have how much package .

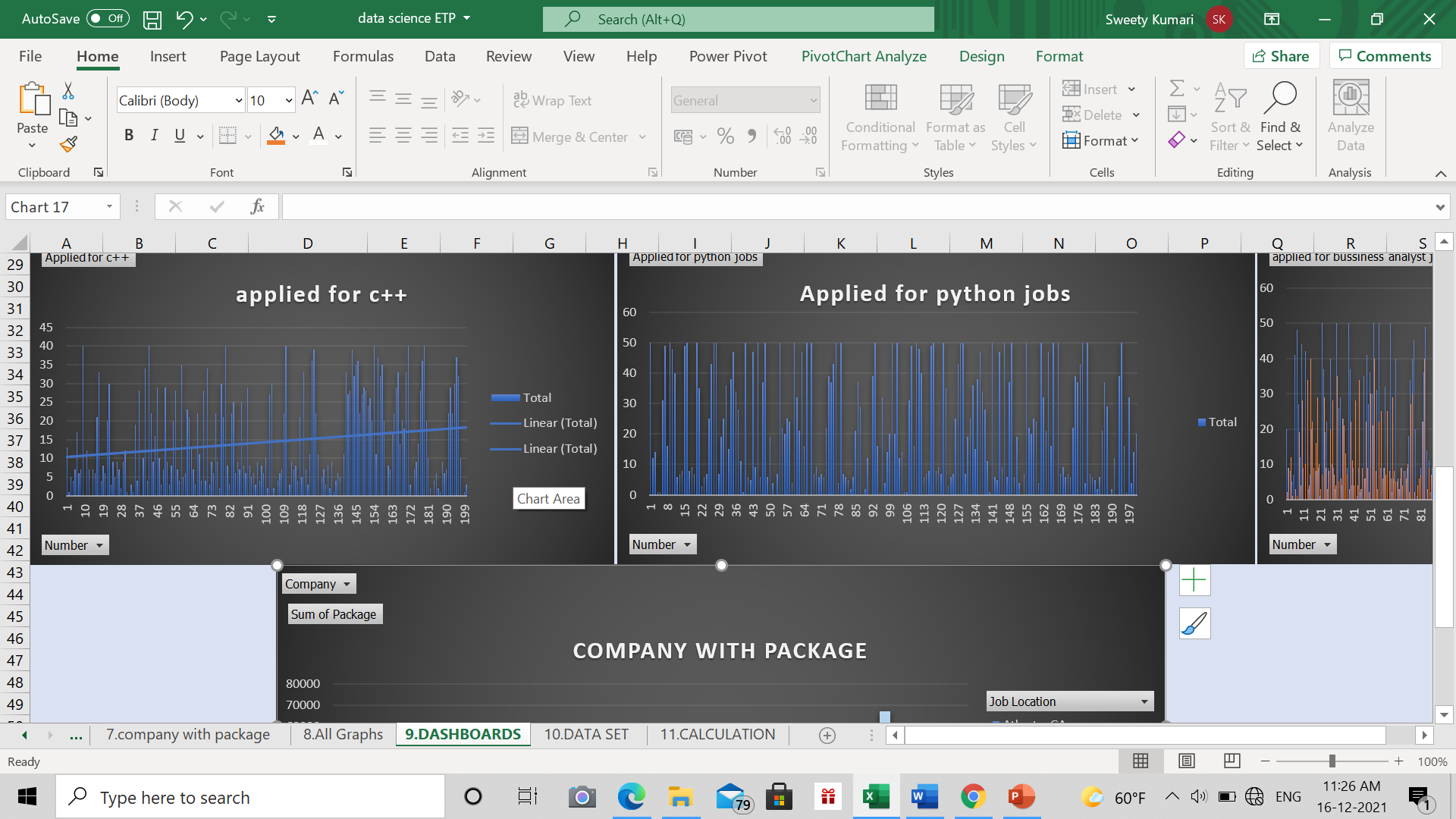
**Results**

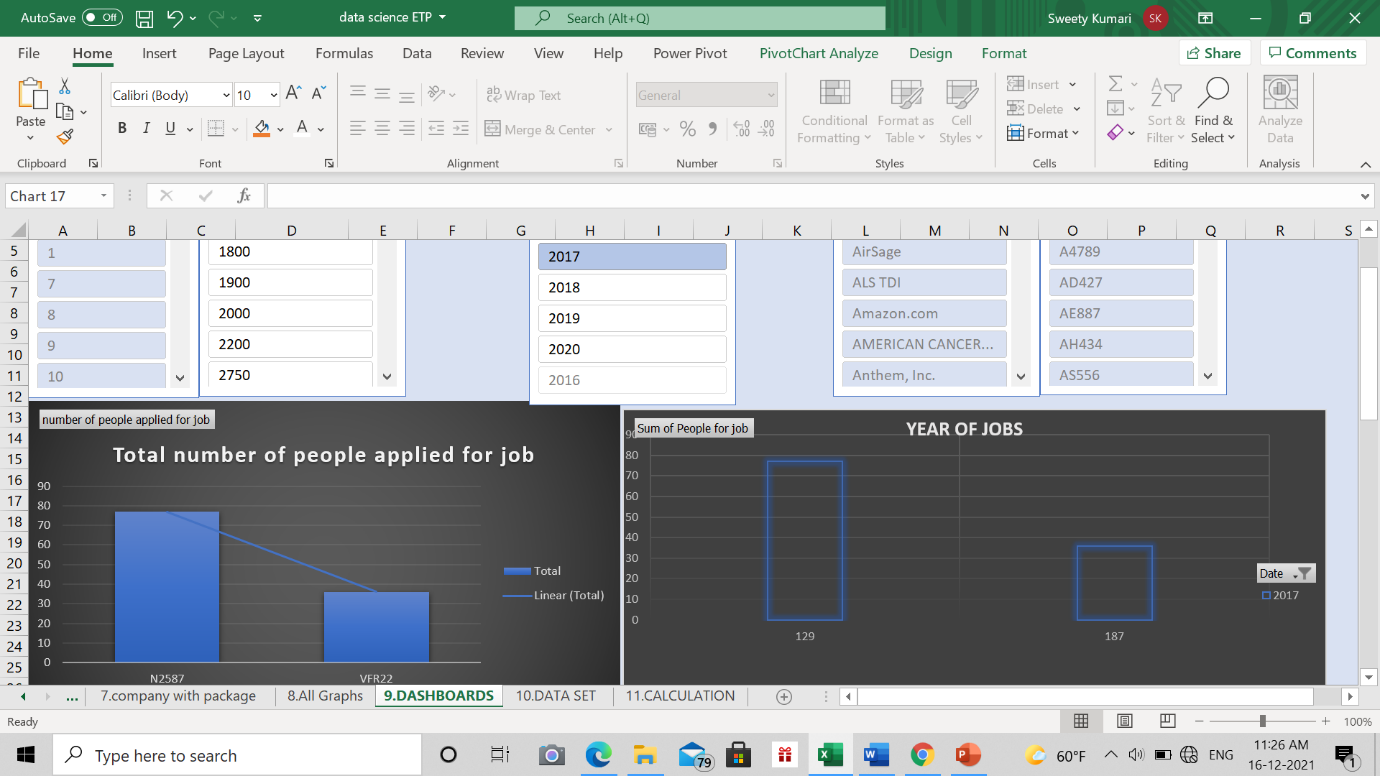
****

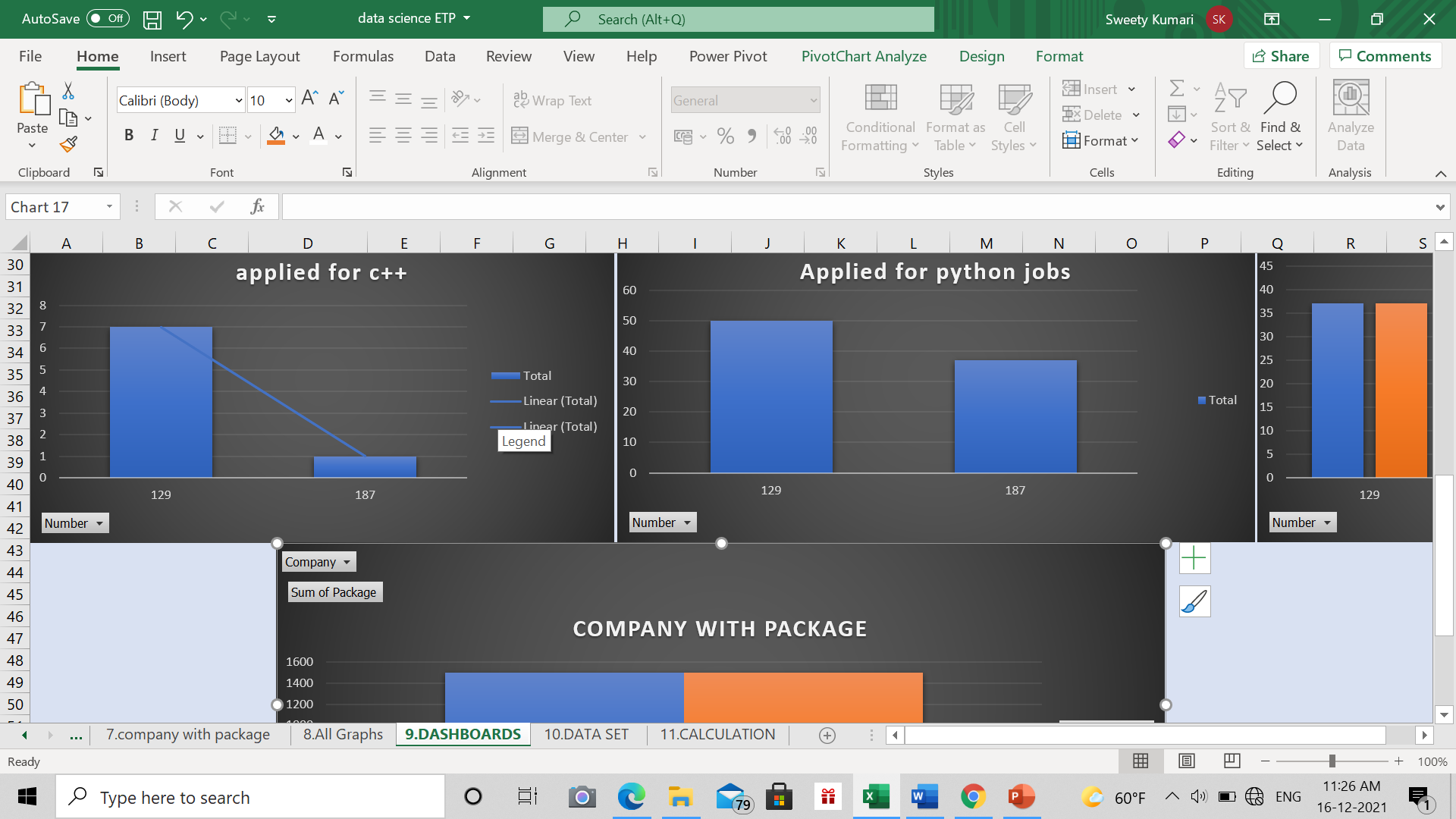
**Visualization**

****

**FINAL DASHBOARD** 



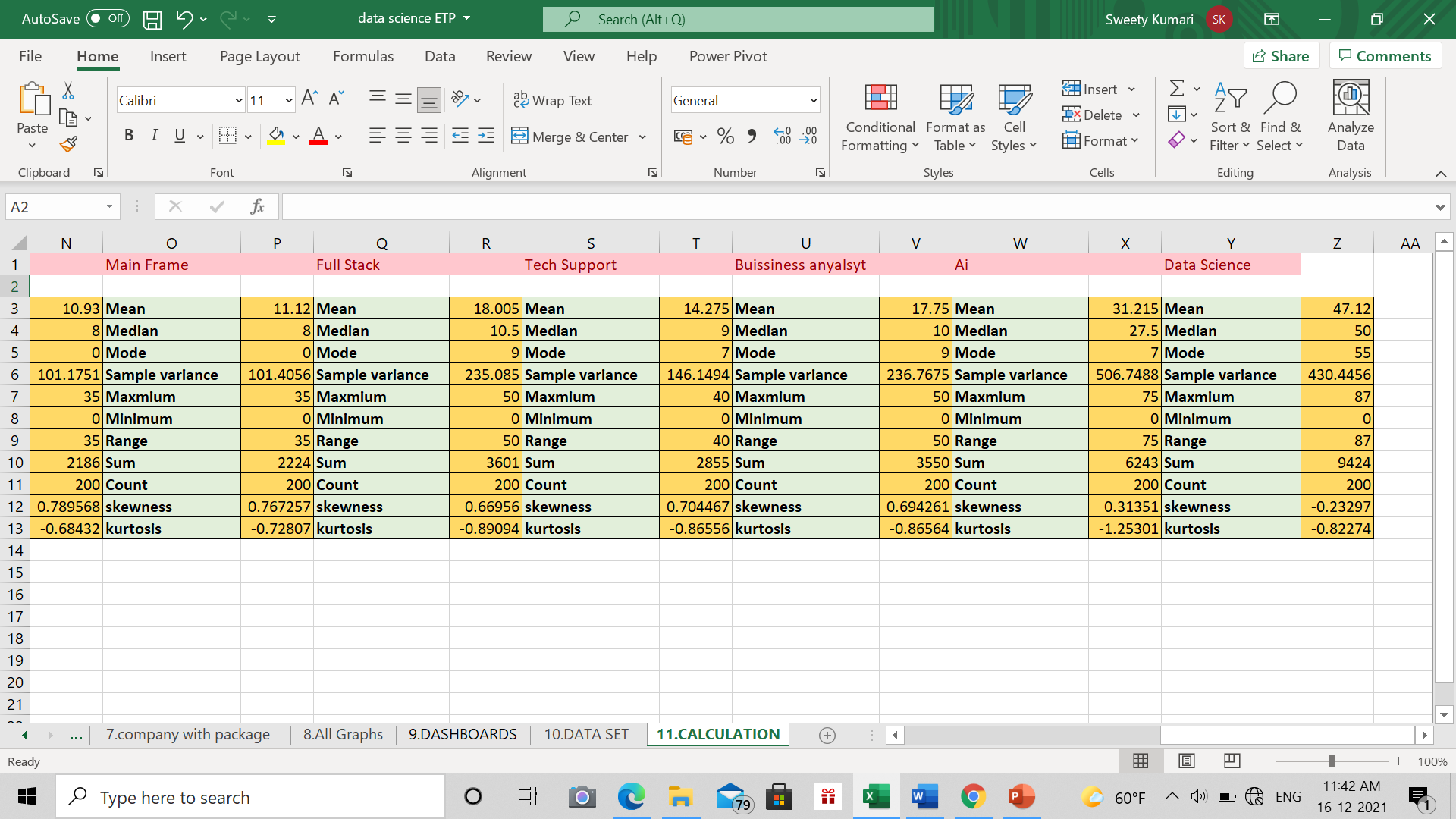


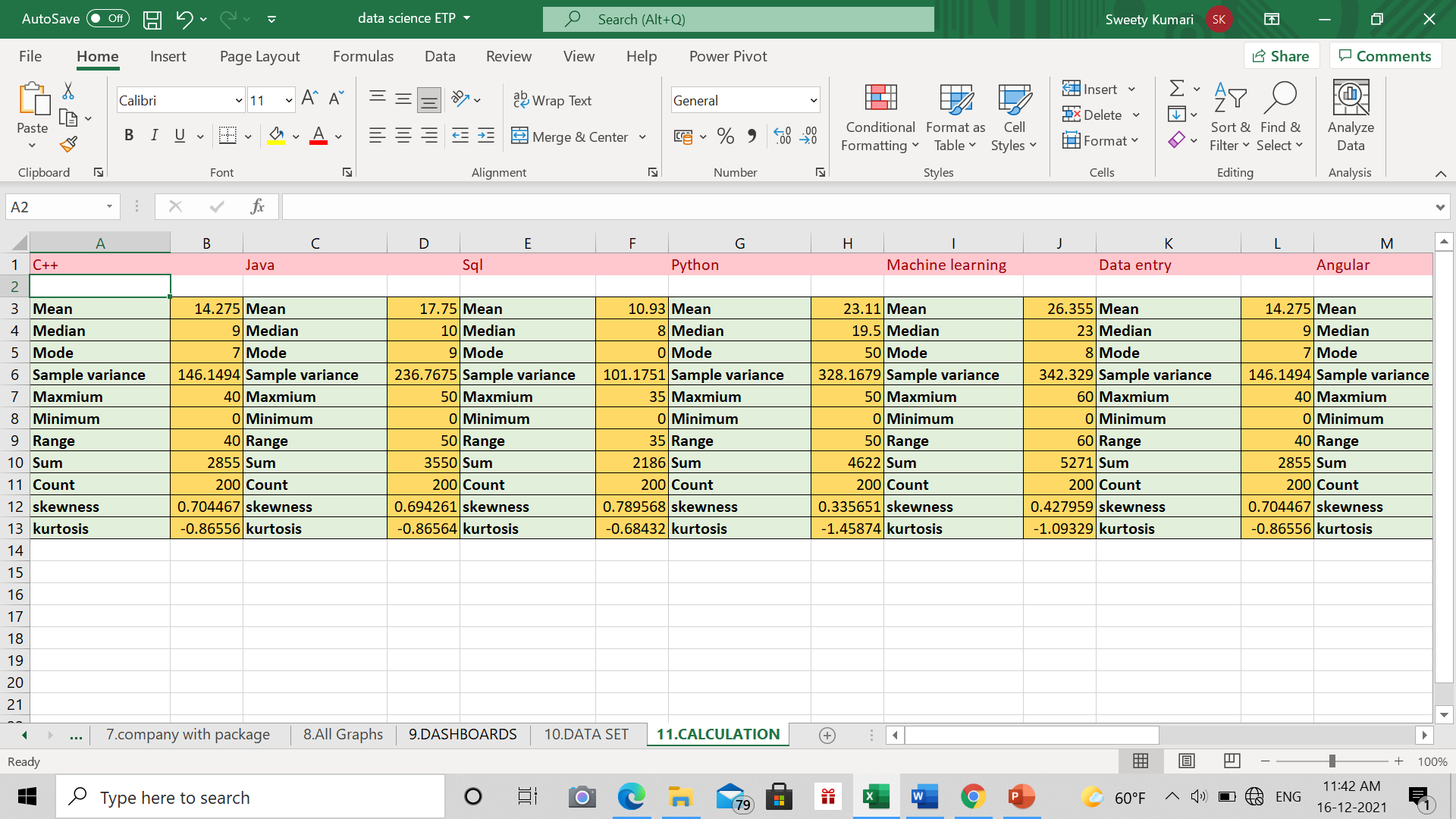
****

**Dashboard Will be giving all the brief details of the of work done on the project.**

**CALCULATION**

Here we analyze the mathematics calculation like median and mode etc. using formulas from the data set “job market in us”.





**Reference and Bibliography**

* [www.youtube.com](http://www.youtube.com)
* [www.kaggle.com](http://www.kaggle.com)
* [www.google.com](http://www.google.com)
* [www.github.com](http://www.github.com)
* [www.Wekipiedia.com](http://www.Wekipiedia.com)
* Microsoft Excel 2016 Bible: The Comprehensive Tutorial Resource by John Walkenbach, Wiley