Contents

[1. Introduction: 1](#_Toc179292678)

[2. Data Collection: 2](#_Toc179292679)

[3. Process: 2](#_Toc179292680)

[4. Methodology: 3](#_Toc179292681)

[5. Experiment: 3](#_Toc179292682)

[5.1. Department vs Annual salary: 3](#_Toc179292683)

[5.2. Country vs Annual Salary: 3](#_Toc179292684)

[5.3. Gender vs Annual Salary: 4](#_Toc179292685)

[5.4. Ethnicity vs Annual salary: 4](#_Toc179292686)

[5.5. Business Unit vs Gender wise Bonus: 4](#_Toc179292687)

[5.6. Department vs Female age: 4](#_Toc179292688)

[6. Discussion: 4](#_Toc179292689)

[7. Conclusion: 4](#_Toc179292690)

[Reference: 5](#_Toc179292691)

# Introduction:

Data analysis, the process of systematically collecting, cleaning, transforming, describing, modeling, and interpreting data, generally employing statistical techniques. Data analysis is an important part of both scientific research and business, where demand has grown in recent years for data-driven decision making.

# 2. Data Collection:

Datasheets are collections of information. Generally, data and datasets are themselves collected to help answer questions, make decisions, or otherwise inform reasoning. The rise of information technology, has lead to the generation of vast amount of data of many kinds, such as text, picture, videos, personal information, account data, and metadata, the last of which provide information about other data. It is common for apps and websites to collect data about how their product are used or about the people using their platform.

# 3. Process:

For data to be analyzed, it must first be collected and stored. Raw data must be processed into a format that can be used for analysis and be cleaned to that errors and inconsistencies are minimized. Data can be stored in many ways, but one of the most useful is in a database. The most familiar kind of database is the relational database, which stores data in tables with rows that represent records and columns that represent fields.

# 4. Methodology:

There are various method for data analysis content analysis, regression analysis, cluster analysis, discourse analysis, factor analysis, grounded theory, statistics, and so on. Among them we can use any type of method for our data to be analysis.

# 5. Experiment:

For data analysis, we use a datasheet which is taken from a website. In this datasheet there show the Employee datasheet from year 1992 to 2021. In the column there are different fields, they are full name, job title, department, business unit, gender, ethnicity, age, hire date, annual salary, bonus %, country, city, exit date and in the row have the value of these fields. For analysis this datasheet we make chart for acknowledge and represent data.

## 5.1. Department vs Annual salary:

A type of chart which is make with the help of the data of department and sum of the annual salary.

## 5.2. Country vs Annual Salary:

It is a pie chart making with Country and sum of the annual salary data of the datasheet.

## 5.3. Gender vs Annual Salary:

A bar chart is made with the help of gender data and sum of the annual salary.

## 5.4. Ethnicity vs Annual salary:

A type of chart which show connection between ethnicity and the sum of the annual salary.

## 5.5. Business Unit vs Gender wise Bonus:

This is a line chart make with Business Unit and sum of gender wise bonus.

## 5.6. Department vs Female age:

A type of chart are drawn with the help of department data and count of female age data which ia available in the datasheet.

# 6. Discussion:

Now we can easily find value of different field which we want to check in this datasheet. So, finding data in this datasheet become easier through connection of chart with one another.

# 7. Conclusion:

Data analysis helps companies evaluate their competitors’ performance, price point, marketing methods, social media reach and so on. It makes it easier to extract information from unstructured data, assisting companies in process optimization, opportunity identification.

# Reference:

Google, Youtube