DATA VISUALIZATION PROJECT REPORT

(Project Semester January-May 2023)

Employee Production Analysis Dashboard

Submitted by

Routhu Siddhartha

Registration No: - 12010599

Programme and Section: - CSE – KM009

Course Code: - CSE233

Under the Guidance of

Ms. Ashu (23631)

Discipline of CSE/IT

Lovely School of Computer Science & Engineering

Lovely Professional University, Phagwara



CERTIFICATE

This is to certify that Routhu Siddhartha bearing Registration no. 12010599 has

completed CSE233 project titled, "Employee Production Analysis Dashboard"

under my guidance and supervision. To the best of my knowledge, the present

work is the result of his/her original development, effort, and study.

Signature and Name of the Supervisor

Designation of the Supervisor

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab.

Date: 18TH April 2023

2

DECLARATION

I, Routhu Siddhartha, student of Bachelor of Technology 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.

Date: 18th April 2023 Signature

Registration No. 12010599 Routhu Siddhartha

Acknowledgement

Place: Phagwara

Date: 18th April 2023

I would like to sincerely thank my manager Ms. Ashu for her significant role behind the accomplishment of the assignment. I have been guided with lots of her valuable suggestions and

experience throughout the process of completion of the assignment.

I would also like to express my gratitude to my peers, without their support and cooperation this

assignment could not have been accomplished. Finally, I would like to thank my parents for their

love and blessings.

Your name: - Routhu Siddhartha

Signature: -

4

Table of Contents

Cover Page	1
Certificate	2
Declaration	3
Acknowledgement	4
Table of content	5
INTRODUCTION	7
BENEFITS OF DASHBOARD	7
Dashboard Data and Format	8
Live Data on Dashboards	8
Features to create Tableau dashboard	8
Data Tables	9
Tableau Charts	9
Filtering the tables	9
Source of Data	10
Formulas used in creating the Dataset	11
ETL Process	12
Use of ETL Process	12
Extraction Process	13
Analysis on Dataset	13
Employee under	14
Production Count	15
DATE	15
AVERAGE TIME TAKEN PER COUNT	15
LOCATION And PROCESS	15

Gender	16
Total Time Taken	17
LIST OF ANALYSIS WITH RESULT	18
1.Employee Production - Avg time VS Sum count	18
2.Process vs Sum of hours taken by the employee	20
3.Each day Production count in process wise.	21
4.KPI for Production count, Avg time/count, Hours and Total time taken	22
5.Gender wise	23
6.Percentage of process wise Production count	24
7.Day wise Employee Production in Tabular chart	25
Insertion Of Filter Dropdown	25
1.FINAL DASHBOARD	26
BIBLIOGRAPHY	30

> INTRODUCTION: -

The core of the dashboard lies in the key metrics required for monitoring. Thus, based on whether the dashboard is for an organization overall or for a department such as sales, finance, human resources, production, etc. the key metrics that are required for display vary.

Further, the key metrics for a dashboard also depend on the role of the recipients (audience). For example, Executive (CEO, CIO, etc.), Operations Manager, Sales Head, Sales Manager, etc. This is since the primary goal of a dashboard in to enable data visualization for decision making.

The success of a dashboard often depends on the metrics that were chosen for monitoring. For example, Key Performance Indicators, Balanced Scorecards and Sales Performance Figures could be the content appropriate in business dashboards.

BENEFITS OF DASHBOARD: -

Dashboards allow managers to monitor the contribution of the various departments in the organization. To monitor the organization's overall performance, dashboards allow you to capture and report specific data points from each of the departments in the organization, providing a snapshot of current performance and a comparison with earlier performance.

Benefits of dashboards include the following –

- > Visual presentation of performance measures.
- ➤ Ability to identify and correct negative trends.
- > Measurement of efficiencies/inefficiencies.
- Ability to generate detailed reports showing new trends.
- ➤ Ability to make more informed decisions based on collected data.
- ➤ Alignment of strategies and organizational goals.
- Instant visibility of all systems in total.

- Quick identification of data outliers and correlations.
- ➤ Time saving with the comprehensive data visualization as compared to running multiple reports.

Dashboard Data and Format: -

The data required for a dashboard depends on its category. The premise for the data is that it should be relevant, error-free, up to date and live if required. The data can possibly be from various and different sources and formats (Spreadsheets, Text Files, Web Pages, Organizational Database, etc.).

You can create a dashboard in Excel using various features that help you make data visualization prominent, which is the main characteristic of any dashboard. You can show data in tables with conditional formatting to highlight the good and bad results, you can summarize the data in charts and PivotTables, you can add interactive controls, and you can define and manage KPIs and so on.

Live Data on Dashboards: -

As discussed earlier in this chapter, data warehousing and online analytical processing (OLAP) is making it possible to refresh the dynamic dashboards instantly with live data. It is also making those who design the dashboards be independent of the organization's IT department for obtaining data.

Thus, the dashboards have become the most sought-after medium from top management to a regular user.

Features to create Tableau dashboard: -

You can create a dashboard in Tableau using various features that help you make data visualization prominent, which is the main characteristic of any dashboard. You can show data in tables with formatting the filters to highlight the good and bad results, you can summarize the data in charts, graphs, and PivotTables, you can add interactive controls, and you can define and manage KPIs and so on.

In this chapter, you will get to know the most important Tableau features that come handy when you are creating a dashboard. These features help you arrive at the dashboard elements that simplify complex data and provide visual impact on the status or performance in real time

Data Tables

The most important component of any dashboard is its data. The data can be from a single source or multiple sources. The data might be limited or might span several rows.

Excel tables are well suited to get the data into the workbook, in which you want to create the dashboard. There are several ways to import data into Excel, by establishing connections to various sources. This makes it possible to refresh the data in your workbook whenever the source data gets updated.

You can name the Excel tables and use those names for referring your data in the dashboard. This would be easier than referring the range of data with cell references. These Excel tables are your working tables that contain the raw data.

Tableau Charts: -

Tableau charts are the most widely used data visualization components for dashboards. You can get the audience view the data patterns, comparisons, and trends in data sets of any size strikingly adding color and styles.

Tableau has several built-in chart types such as line, bar, column, bubble, pie, area, Funnel, Bump, Dual axis charts and so on.

Filtering the tables

When you have large data sets and you would like to summarize the results dynamically showing various facets of the analysis results, Tableau Filtering come handy to include in your dashboard. You can use either the data tables or the more powerful data tables in the data model to create Filtering.

The main differences between the two approaches are –

Data Tables

- Data from more than one table can be used to create Filtering, defining relationships between the tables.
- Can handle huge data sets with thousands of rows of data with memory optimization and decreased file size.

> Source of Data

A dataset is a collection of data within a database.

Typically, datasets take on a tabular format consisting of rows and columns. Each column represents a specific variable, while each row corresponds to a specific value. Some datasets consisting of unstructured data are non-tabular, meaning they do not fit the traditional row-column format.

As I have prepared the dataset by my own, I have taken Employee production analysis dashboard by labelling the contents as follows

Employee Under, Production Count, Date, Average time per count, Location, Gender, Process, Total time taken (Hours, Days, Months, Years, mmm-dd)

I have prepared the columns by applying Randbetween function for getting random values from top to bottom

So, that a random number will dice on cell. Then you need to drag the Doctor's symbol (Big Plus Symbol) to the down of the cell until where we want to apply the dataset.

After that the numbers formulated in the cell will change every time, we use any cell for operation. so, copy the data from that cell and paste again by the option paste values in the same place so, that the data in the cells will not change.

> Formulas used in creating the Dataset

- o = "Name"&RANDBETWEEN(2235,2245)
- \circ =RANDBETWEEN(45,65)
- o =RANDBETWEEN(TODAY(), TODAY()+5)
- o =RANDBETWEEN(5,7)
- o =VLOOKUP(A3,\$N\$4:\$O\$15,2,0)
- o =VLOOKUP([@[Employee under]],\$O\$4:\$R\$15,4,FALSE)
- =VLOOKUP(A3,Emp_Process,3,0)
- \circ Total Time taken in Hours = (total time taken)/60; used : =G3/60
- Day of the production, formula is =TEXT(C3,"DDDD")
- o For getting the Particular Month of the data is =TEXT(D3,"mmm")
- For getting the Year of the data is =TEXT(D3,"yyyy")
- \circ For getting the format of data in mmm-dd is =TEXT(D3,\$M\$2)

> ETL Process

ETL, which stands for extract, transform and load, is a data integration process that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse or other target system.

As the databases grew in popularity in the 1970s, ETL was introduced as a process for integrating and loading data for computation and analysis, eventually becoming the primary method to process data for data warehousing projects.

ETL provides the foundation for data analytics and machine learning workstreams. Through a series of business rules, ETL cleanses and organizes data in a way which addresses specific business intelligence needs, like monthly reporting, but it can also tackle more advanced analytics, which can improve back-end processes or end user experiences. ETL is often used by an organization to:

- Extract data from legacy systems
- o Cleanse the data to improve data quality and establish consistency
- Load data into a target database

The data is loaded in the DW system in the form of dimension and fact tables.

ETL combines all the three-database function into one tool to fetch data from one database and place it into another database.

Use Of ETL Process

ETL is used to integrate the data with the help of three steps Extract, Transform, and Load, and it is used to blend the data from multiple sources. It is often used to build a data warehouse.

In the ETL process, data is extracted from the source system and convert into a format that can be examined and stored into a data warehouse or any other system.

Extraction Process: -

• Extract is the process of fetching (reading) the information from the database. At this stage, data is collected from multiple or different types of sources.

• A staging area is required during ETL load. There are various reasons why staging area is required.

The source systems are only available for specific period of time to extract data. This
period is less than the total data-load time. Therefore, staging area allows you to extract
the data from the source system and keeps it in the staging area before the time slot
ends.

> Analysis on Dataset

A dataset is a collection of data within a database.

Typically, datasets take on a tabular format consisting of rows and columns. Each column represents a specific variable, while each row corresponds to a specific value. Some datasets consisting of unstructured data are non-tabular, meaning they don't fit the traditional row-column format.

As I have prepared the dataset by my own, I have taken Employee production analysis dashboard by labelling the contents as follows

- > Employee Under
- Production Count
- > Date
- > Average time per count

> Location
Gender
> Process
> Total time taken
o Hours
o Days
o Months
o Years
o mmm-dd
1.Employee under: -
In the first column of the table, I mentioned the details of the Co-Ordinator's ID for the following workers for having the better result on getting initiation for individual ID.
I have taken this column by applying Randbetween function for getting random values from top to bottom
= "Name"&RANDBETWEEN(2235,2245)
So, that a random number will dice on cell. Then you need to drag the Doctor's symbol (Big Plus
Symbol) to the down of the cell until where we want to apply the dataset.

After that the numbers formulated in the cell will change every time, we use any cell for operation. so, copy the data from that cell and paste again in the same place so, that the data in the cells will not change.

2.Production Count: -

It shows the details of the production count which was made by the Employee.

I have apply the formula, =RANDBETWEEN(45,65) and do the same procedure which we done in previous step.

3.DATE: -

I have taken particular time to implement for the particular result, for the same I have applied the formula,

=RANDBETWEEN(TODAY(), TODAY()+5) and change the format of the cell data in the particular format which you want. And continue the process of dragging down the doctors plus symbol until we want to take.

4.AVERAGE TIME TAKEN PER COUNT: -

As usual the average time taken to complete the work for individual worker would be approximately 5 hrs to 7 hrs. so, I have applied the formula

=RANDBETWEEN(5,7) to get random number as employee time taken on particular day.

5.LOCATION And PROCESS: - Before this, we need to apply a pivot table for the table which we created till now, by selecting only Employee section to get all the IDs of the employee under section.

Then, do copy the data of the pivot table and paste in some other cell and do delete the pivot table. The new table name "Emp-Process". We are doing this procedure for getting the list of Employee's Under section. Later, next to the pasted values add some location names as you wish.

And now come to our previous dataset to add the same locations over all the datasets. For the same, I apply the formula of vlookup as =VLOOKUP(A3,\$N\$4:\$O\$15,2,0) to get the designated location for our dataset and drag down the doctor's plus to whole over our dataset.

Now, from the same table Emp-Process, we add one more column name process and add the details of the process done by the Employee's Under section.

Apply the same, vlookup formula on our dataset as =VLOOKUP(A3,Emp_Process,3,0) to get the data of the process done also without any problem

6.Gender

The new table name "Gender". we need to apply a pivot table for the particular table which we created till now, by selecting only Employee section to get all the IDs of the employee under section.

Then, do copy the data of the pivot table and paste in some other cell and do delete the pivot table. The new table name "Emp-Process". We are doing this procedure for getting the list of Employee's Under section. Later on next to the pasted values add some location names as you wish.

And now come to our previous dataset to add the same locations over all the dataset. For the same, I apply the formula of vlookup as =VLOOKUP([@[Employee

under]],\$O\$4:\$R\$15,4,FALSE) to get the designated gender for our dataset and drag down the doctor's plus to whole over our dataset.

Now, from the same table Emp-Process, we add one more column name gender and add the details of the process done by the Employee's Under section.

Apply the same, vlookup formula on our dataset as =VLOOKUP([@[Employee under]],\$O\$4:\$R\$15,4,FALSE) to get the data of the process done also without any problem

7. Total Time Taken: -

Here, total time taken by the employee for the particular product will come by multiplying the avg. time count by Production count.

As, I have applied in the section with the formula =D3*B3 for getting the same and drag the pointer.

Total Time taken in Hours = (total time taken)/60; example : = G3/60

Particular Day of the production, formula is =TEXT(C3,"DDDD")

For getting the Particular Month of the data is =TEXT(D3,"mmm")

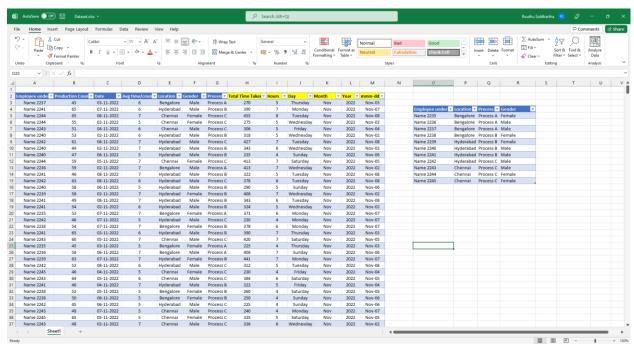
For getting the particular Year of the data is =TEXT(D3,"yyyy")

For getting the format of data in mmm-dd is =TEXT(D3,\$M\$2)

> LIST OF ANALYSIS WITH RESULT :-

1. Dataset

My dataset



The Data continuous with information for 1000 lines. As it is good to take big dataset for better visualization.

1. Employee Production - Avg time VS Sum count

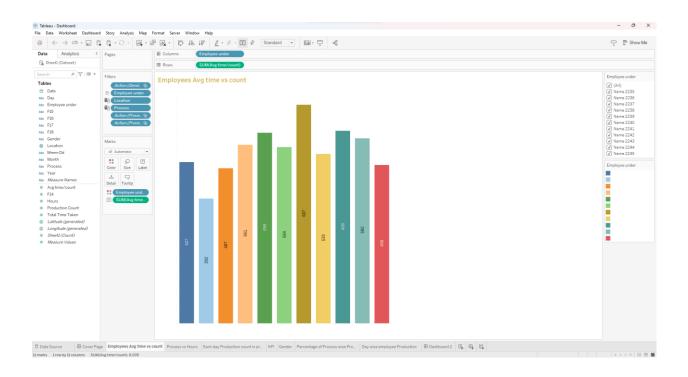
First, we need to drag and drop the dimensions and measures to the rows and columns for the data which was needed. Second method for the same is to double click on the measures first and then give a double click on the dimension so that tableau automatically identifies the data and gives the desired and relevant graph for that dimensions and measures which we imported in rows and columns

Here, the dimension for this section Employee production – Avg time vs Count from the data is Names of the employees as employee under along with the measure Avg time per count. Double click on both of them, tableau automatically will detect the data type and accordingly it give the bar chart as automatically. Now, drag the Names of the employees which is Employees under to the filter panel and drop it in the color tab option. So, that this will show the different colors on the interface of the Bar chart according to the Employee. As we can also customize the color from the edit color option which is in the filter panel of color tab. Change it accordingly.

In the same way drag the Avg time/count section in the filter option and drop it in the label section to show the data of the avg time/count for that employee on the rectangular bar.

Last step is just given right click on Employee under and select on show filters. So, that we can manipulate the data visualization for employee as well.

This gave me the final Bar chart for this data Visualization to keep in the dashboard.

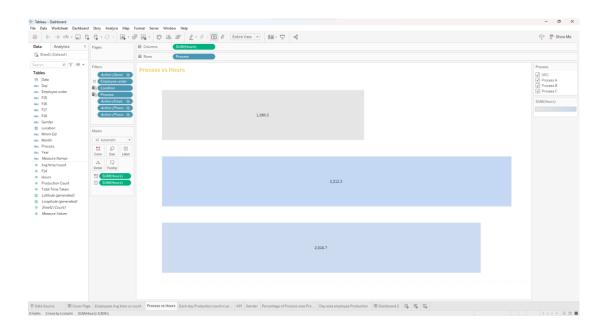


In this sheet it shows the average of the production count and the sum of production count done by the employee. We created this table to get visualize the clear-cut data of the employee for the same along with the bar chart which consists of bar graph

2. Process vs Sum of hours taken by the employee

Same as before, we need to drag and drop the dimensions and measures to the rows and columns for the data which was needed. Second method for the same is to double click on the measures first and then give a double click on the dimension so that tableau automatically identifies the data and gives the desired and relevant graph for that dimensions and measures which we imported in rows and columns

Here, the dimension for this section Process vs Sum of Hours taken for single process by the employees. Double click on both, tableau automatically will detect the data type and accordingly it gives the bar chart as automatically. Click on show me option in the top right corner and select horizontal bar charts as a graph. Now, drag the Sum of hours to the filter panel and drop it in the color tab option. So, that this will show the different colors on the interface of the horizontal Bar chart accordingly. As we can also customize the color from the edit color option which is in the filter panel of color tab. Change it accordingly.



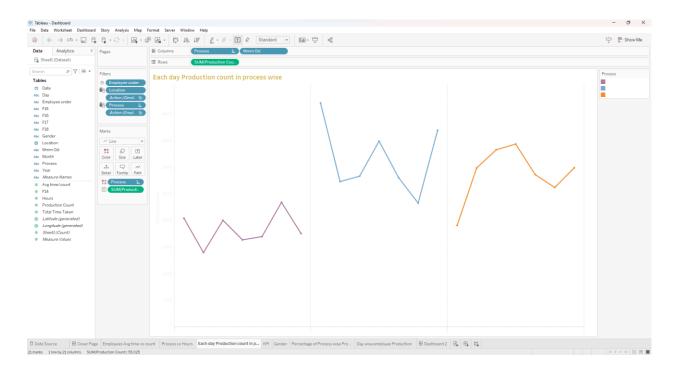
In the same way drag the sum of hours section again in the filter option and drop it in the label section to show the data of the time taken in hours for that employee in the rectangular bar.

Last step is just to give right click on Process and select on show filters. So, that we can manipulate the data visualization for different processes as well. This gives me the final Bar chart for this data Visualization to keep in the dashboard.

3. Each day Production count in process wise.

Same as previous one, we need to drag and drop the dimensions and measures to the rows and columns for the data which was needed. Second method for the same is to double click on the measures first and then give a double click on the dimension so that tableau automatically identifies the data and gives the desired and relevant graph for that dimensions and measures which we imported in rows and columns

Here, the dimension for this section is Process in days(mmm-dd) vs Production count for single count by the employees. Double click on both, tableau automatically will detect the data type and accordingly it gives the chart as automatically. Change the chart to line graphical chart. Click on show me option in the top right corner and select line chart. Now, drag the process to the filter panel and drop it in the color tab option. So, that this will show the different colors on the interface of the line graph accordingly. As we can also customize the color from the edit color option which is in the filter panel of color tab. Change it accordingly

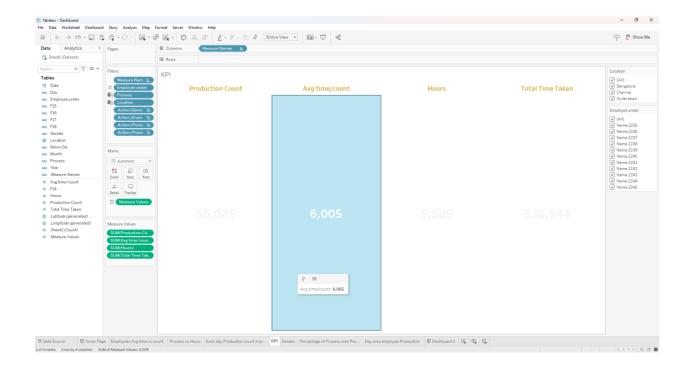


In the same way drag the Production count section again in the filter option and drop it in the label section to show the data of the time taken in production count for that employee on the line graph. This gives me the final line chart for this data Visualization to keep in the dashboard

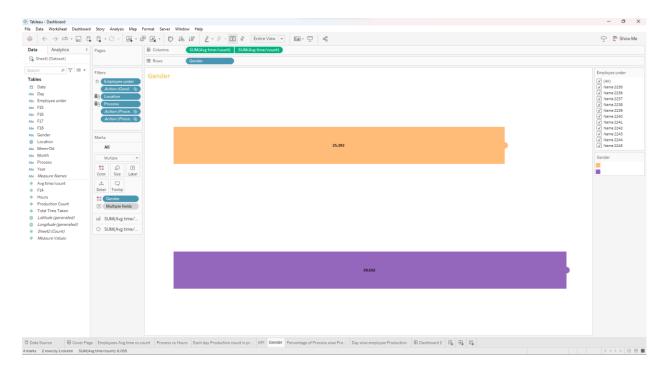
4. KPI for Production count, Avg time/count, Hours and Total time taken

This case is little different from other. Double click on the Measure names from the list. Automatically the data will come in the form of small matrices. Change the standard view to entire view. Add the data columns which you want to show in the KPI.

Drag and drop the Production count, Avg time/count, Hours and total time taken sections to the measures panel under filter panel. Now, change the data colors accordingly.



5. Gender wise: -



This chart is known as dual axis chart. Over here it is known as lollipop chart.

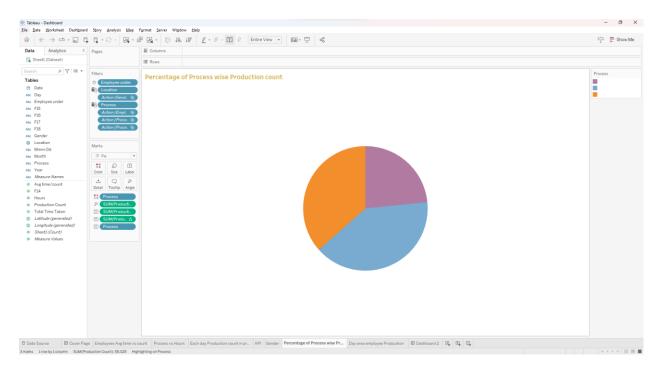
First double click on gender as it will import in rows. Then double click on Avg time/count to fix it in column. Again, bring the same Avg time/count section to the column for the second time

and make it dual axis for both by selecting in right click on the column section. Now change the chart for both like horizontal bar graph and circular chart. Now, right click and select the option synchronized axis option to make both the charts to visible on a single axis as a single graph.

Filter the gender with customized colors by drag and drop on colors panel in filters. And then drag and drop the multiple values to the label section to show the avg time/count on the bar rectangle.

6. Percentage of process wise Production count

This visualization is shown by the pie chart. First drag and drop the process section in the colors panel in the filter without adding anything in the rows and columns. Select the type of chart from automatic to pie chart. This will give the pie chart with same angle for the process. We need to change the angle of the pie chart by applying different category in this field. So, drag and drop the production count to the Angle panel which is there in the filter panel option. Now, apply changes according to the processes to identify easily. And label the pie chart with production count and process name

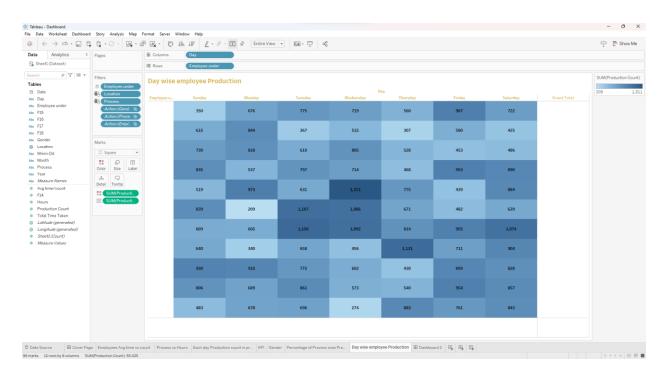


This will give the best visualized pie chart with the better implementation without inserting the type of data in the rows and columns.

7. Day wise Employee Production in Tabular chart

This data is visualized in the form of tabular chart of day wise employee production. Drag and drop the Employee under section in rows and days in columns. As they both are dimensions, automatically the data visible in the form of tabular format. Now, take Employee production count and drop it in the label panel of filtration. It will show the data in the table accordingly. Then go to analysis and click on totals > totals for rows and columns. Grand totals are now added to this tabular chart.

Now drag and drop production count in the color section also to get dark color for high rate of work and light color on least performed production of an employee. There the numbers of the data will get colored. Click on show me option above and select highlight map, to convert the table into color highlighted tabular chart which looks better. Swap the rows as accordingly to get final chart and select entire view.



Insertion Of Filter Dropdown: -Filter dropdown provide buttons that you can click to filter tables, or Pivot tables. In addition to quick filtering, it also indicates the current filtering state, which makes it easy to understand what exactly is currently displayed.

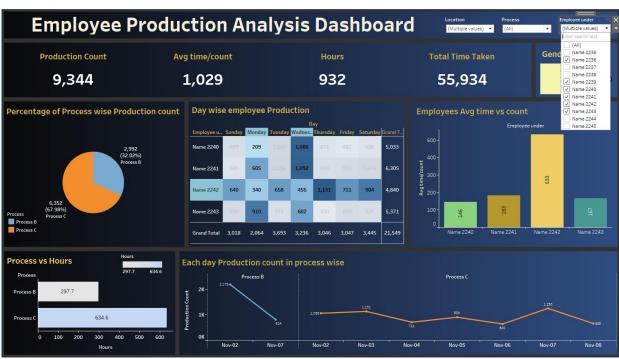
Create a dropdown to filter data: -

- Right click on the filter section and select the floating option.
- Paste the section on the position
- Again, right click on the filter box, then select the option multiple dropdown values option.
- This will show the data filtration box in simple and compact size dropdown menu option

1. FINAL DASHBOARD







Dashboards are a powerful way to visualize data. They have become a popular business tool over the years. Analyzing data is easier than ever with the help of an tableau Dashboard. The tableau Dashboard can be very captivating and helps users get an insight on the data just by taking a glance at it. Whether you want to customize your reports or keep track of metrics, these dashboards get the job done effortlessly.

The tableau dashboard provides an overview of metrics and other data points in one place. In simple terms, dashboards are visual representations of data. They mostly consist of charts and graphs, thereby grabbing the user's attention. Looking at raw Excel data can be boring. Creating a Dashboard in Excel can help you interpret the data by giving an advanced level overview of the same.

Customize the Chart Accordingly

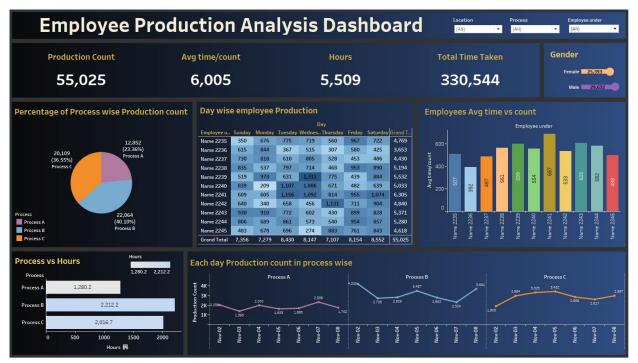
- To make your charts appealing and attractive, we can change the graph's colors, add text, or give more information about it, etc.
- To do this, right click on the chart and use the different customization options available in the toolbar.
- You can also select the format option right next to the chart.
- This will open a drop-down box panel consisting of various Chart Elements customization.

An tableau dashboard includes numerous charts and graphs. So, go ahead and add more visual elements to your Dashboard, as per your choice.

You can use the filter option to filter out the data in your database and form appropriate charts. In the following Dashboard.

As you can see, these elements together help us track various metrics and make complex datasets easier to analyze. With this, you have reached the end of the process of building your personalized Dashboard from a dataset.

A dashboard is a visual display of all your data. While it can be used in all kinds of different ways, its primary intention is to provide information at-a-glance, such as KPIs.



A dashboard usually sits on its own page and receives information from a linked database. In many cases it is configurable, allowing you the ability to choose which data you want to see and whether you want to include charts or graphs to visualize the numbers.

Dashboards offer a method of consolidating company data into one unified location with secure data storage. Dashboards are designed to offer a comprehensive overview of company performance and do so using data visualization tools like charts and graphs.

Dashboards are built to provide quick insights into some of the most important business processes. Dashboards work best if the information they contain is to the point and instantly visible. The dashboard-building process begins with determining its purpose and key performance indicators involved.

The Summary dashboard gives you a broad overview of the customer service experience in your organization. It uses AI to provide insights into topics that are generating the highest volume and which topics that are emerging with the highest rate of change in volume.

BIBLIOGRAPHY

- o https://catalogimages.wiley.com/images/db/pdf/9781118490426.excerpt.pdf
- o https://www.ibm.com/in-en/cloud/learn/etl
- $\circ \quad \underline{https://help.tableau.com/current/pro/desktop/en-us/dashboards_create.htm}$
- o https://www.upwork.com/resources/how-to-create-a-dashboard-in-tableau