**DESCRIPTION for the Columns in the Data set**

The HR\_Data\_Set is divided into three sub datasets.

1. HRIS-1(HR Information system 1)

2. HRIS-2(HR Information system 2)

3. FIS (financial information system)

Columns in HRSI-1

General HR information gathered from the HR information system I

A. Employee ID: Unique employee record number .

B. Position: Role the employee held.

C. DOB: Date of birth of the employee.

D. Gender: Employee gender (Male / Female).

E. Marital Status of the employee: Single / Married / Divorced / Widowed.

F. Date of hire: Date when the employee was hired by the company.

G. Date of termination: If the employee has been terminated, the date on which they were terminated or the last official working day.

H. Termination reason: The reason why the employee left the organization.

I. Employment status: Whether the employee is currently with the company, terminated for a reason, or they left themselves.

Columns in HRSI-2

Performance and Employee satisfaction survey results gathered from the HR information system II

A. Employee ID: Unique employee record number

B. Dept: which department the employee worked in – Admin / Executive Office / IT /Operations / Sales / Software Engineering

C. Manager name: Employee’s manager name

D. Perf Score: Most recent employee performance rating –Exceeded expectations / Met Expectations / Improvement needed. Employees with ‘Improvement needed’ are generally moved towards a performance improvement program in which they are monitored for about 3 – 6 months after which a decision is made whether to continue to employ them let them go depending on the performance during the improvement program.

E. Emp. Sat: Employee satisfaction score – rating given by the employee on how much they are satisfied with their employment with the company. The survey is conducted once a year in a confidential way. A rating of 1 means the employee is highly dissatisfied and a rating of 5 means that the employee is highly satisfied. However, it is noticed that the employees very rarely gave a score of 1 or 2

F. Date of the last perf. Review – Date on which the most recent performance review was conducted

G. Late days: how many days the employee came late to work in the last one-month time

H. Absent days: how many leave days the employee availed in the last one year

Columns in HRSI-2

Employee salary details gathered from the financial information system

A. Employee ID: Unique employee record number

B. Salary: Annual salary of the employee in USD

1. **Preparing the data for Tableau**

Before conducting any analysis in tableau, we need to first connect the data and see if there are any issues in the data set and then clean it up.

**Connecting:**

The data connection steps are given below

1. Connect tableau by selecting Connect 🡪 To a file🡪Microsoft Excel 🡪Select file and open

2. In this case, since all sheets have equal number of rows and the employee ID is available in all sheets, we can do any type of join or data blending or combine using relationships

3. We will combine using relationships

4. Since HRIS-1 is the sheet that contains most information, we will bring it first to the tables section. Tableau will automatically use Employee ID to connect the tables. Close the window and move forward to connect the sheet FIS



5. Final data model should look like this



**2.Data clean up**

In this case, the data is quite clean. We will do a couple of steps to make the data more usable for future analysis. This involves creating grouping the data and creating hierarchies.

* 1. 1. Termination reasons are quite detailed. So, we will group them into some high-level groups. Note that this grouping is subjective and everyone would come up with a different grouping depending on how they would analyze
  2. a. *Career issues:* Career change, found a better job, went for higher paying job
  3. b. *Personal issues:* Did not return from maternity, health, higher education, relocated, sabbatical
  4. c. *Personnel issues:* Absconded, attendance, legal issues, misconduct, performance issues, Unhappy
  5. 2. Positions are also quite many in the dataset. It would be worthwhile to bring in the hierarchy here so that it makes more sense in the analysis. For instance, we can group the positions into senior management, middle management and employees
  6. a. *Senior management:* CEO / CIO/ Directors
  7. b. *Middle management:* Anyone with having ‘Manager’ in their job title
  8. c. *Employees:* everyone else

1. *Employment status:* we will combine ‘terminated for cause’ and ‘voluntarily terminated’ as one group called ‘terminated’.
2. We will rename the original ‘Employment status’ field as ‘Termination type’. This will allow use to nest the two fields employment status and termination type together if we want.

2**. Visualizing HR Data at Acme**

We will prepare the charts and dashboards in the order of expectations laid out. Note that there is no one way in terms of what charts need to be used.

**Exploratory Analysis**

1) Group size and demographics

a. How many employees are currently employed by each department?

b. What are the demographics of our current employees?

i. Age

ii. Gender

iii. Marital Status

2) Salary structure

a. What was the current total salary expense for each department?

b. What is the salary structure for each demography of our current employees?

i. Age

ii. Gender

iii. Marital Status

3) Performance results

a. What was the distribution of employees in terms of their performance?

b. Could we do a deep dive per group?

i. Department

ii. Age

iii. Gender

iv. Marital Status

4) Satisfaction scores

a. How satisfied our employees are?

b. Could we do a deep dive per group?

i. Department

ii. Position

iii. Age

iv. Gender

v. Marital Status

5) Attrition

a. How many employees have left the company in total?

b. What were the main reasons for them to leave?

c. How many of those reasons are voluntary and non-voluntary?

d. Is there some pattern related to employee groups?

i. Department

ii. Age

iii. Gender

iv. Marital Status

**Exploratory Analysis**

1) Group size and demographics

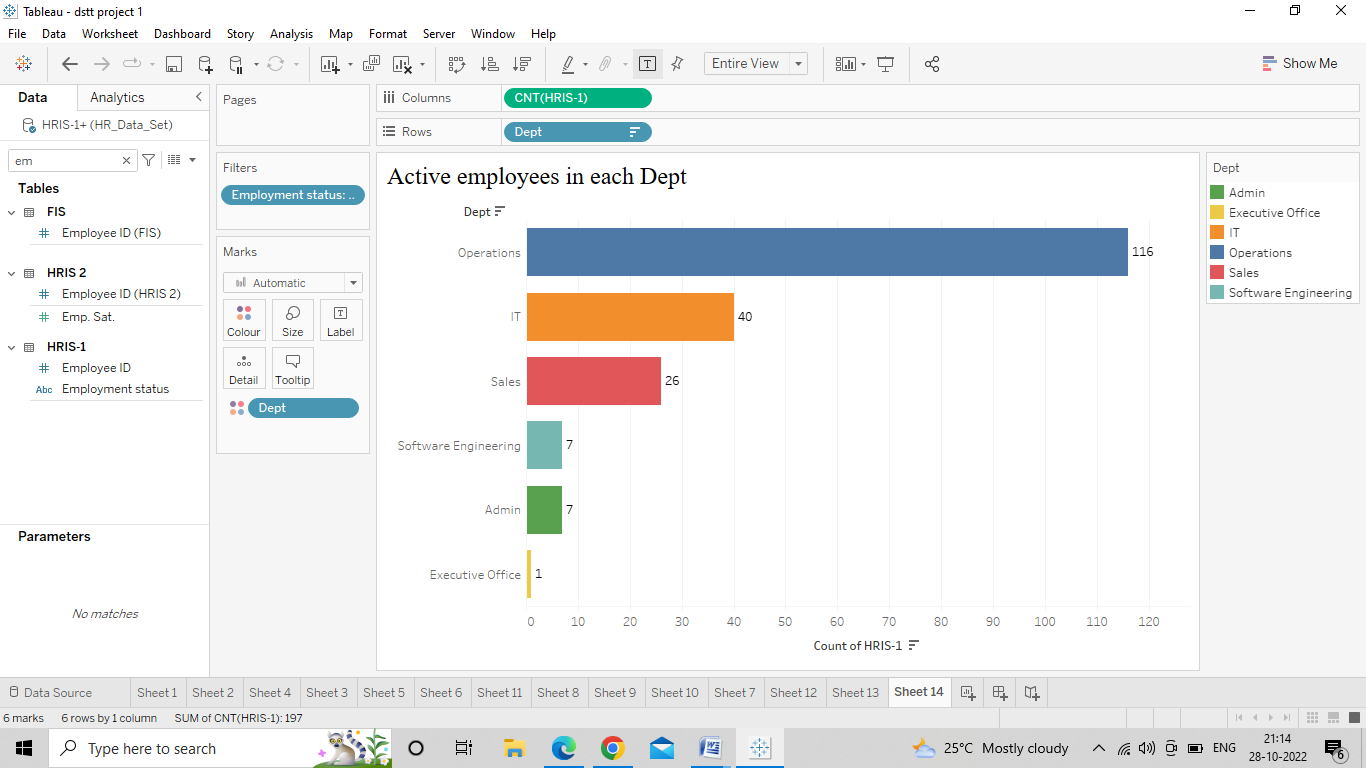
a. How many employees are currently employed by each department?

● Move *Dept* to rows and *HRIS-1 (count)* to columns

● Move *Employment Status* to filter and filter ‘*Active employees’*

● Show text labels

● Sort the bar chart in the descending order



**Insight:**

The largest department is Operations with 116 employees, followed by IT, Sales, Admin, Software engineering and Executive office with 40, 26, 7, 7, and 1 employee respectively.

b. What are the demographics of our current employees?

i. Age

● Create a calculated field called ‘Review Date’ with text ’31-Dec-2020’

● Create a calculated field to compute age (‘Age’) with the following calculation:

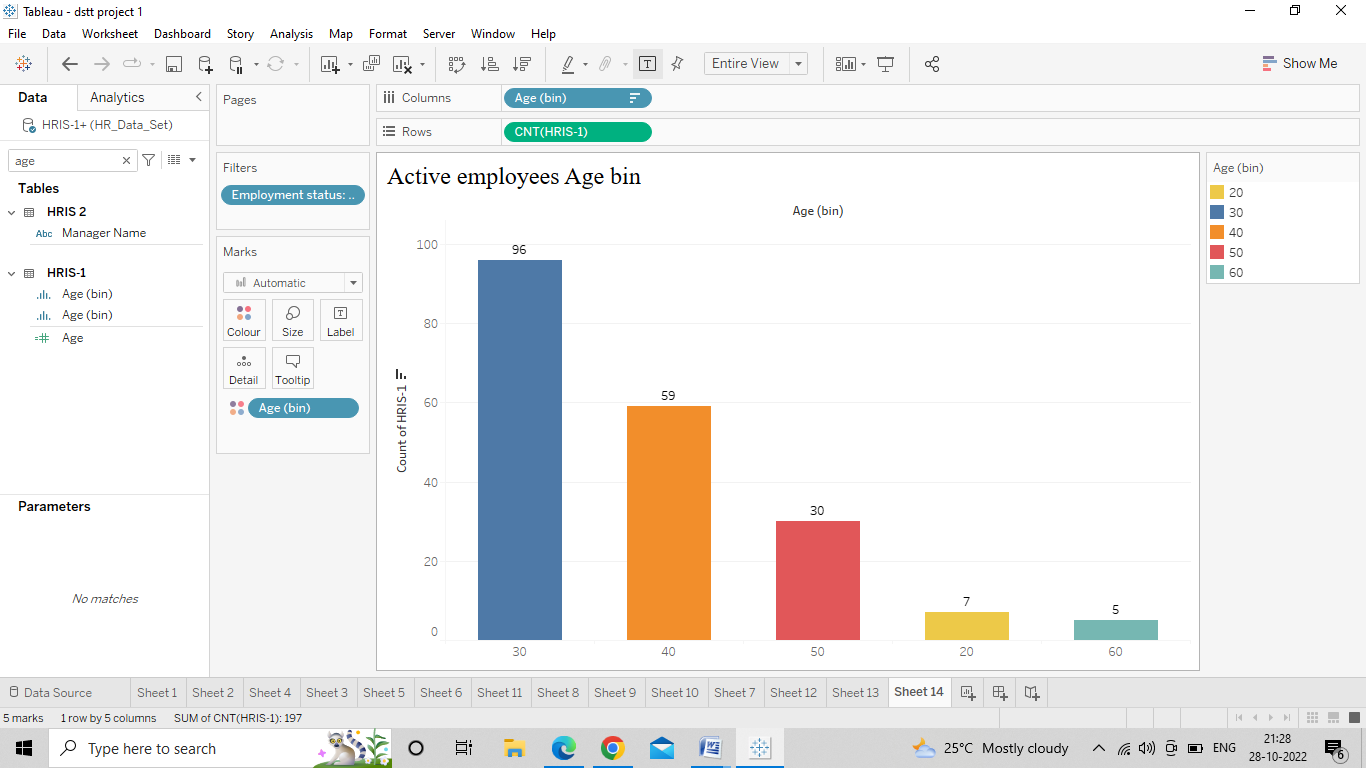
**ROUND(([Review date]-[DOB])/365,0)**

This computes the days between date of birth and the review date, converts it into years and then rounds off to the nearest year.

● Create bins for the field ‘*Age’* with a bin size of 10 years

● Bring ‘*Age (bin)*’ to columns and HRIS-1 (Count) to rows

● Filter for active employees



**Insight:**

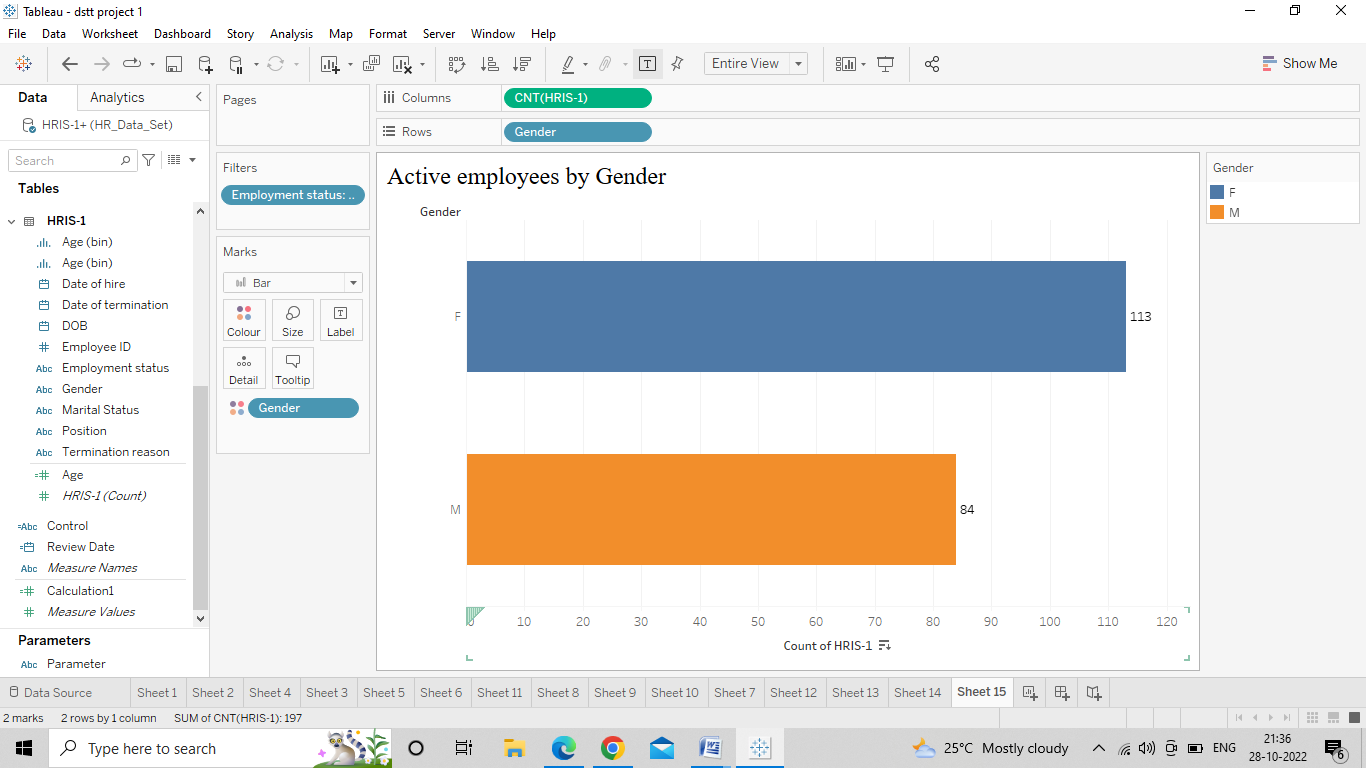
The age group with most number of employees is 30 – 40 years with 96 employees followed by 40 – 50 years, 50 – 60 years, 20 – 30 years and then finally > 60 years with 59, 30, 7 and 5 employees respectively.

3) 2**What are the demographics of the current employees by Gender?**

● Bring Gender to rows and HRIS(Count) to columns

● Filter for active employees

● Enable text labels



**Insight:**

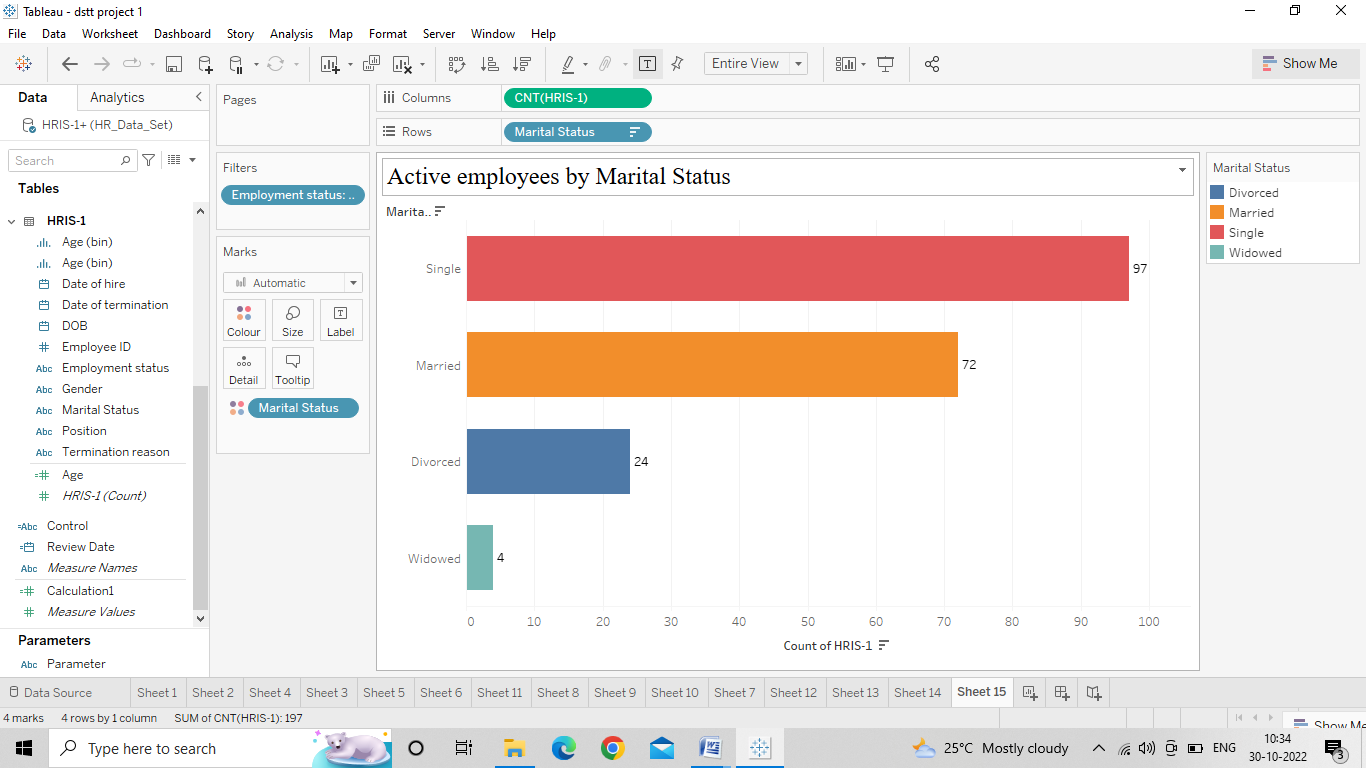
The company has more females than male employees – 113 and 84 respectively

4) **What are the demographics of the current employees by Marital Status?**

● Bring Marital status to rows and HRIS(Count) to columns

● Filter for active employees

● Enable text labels



**Insight:**

Most of the employees are single – 97, followed by married, divorced, and widowed – 72, 24, and 4 respectively.

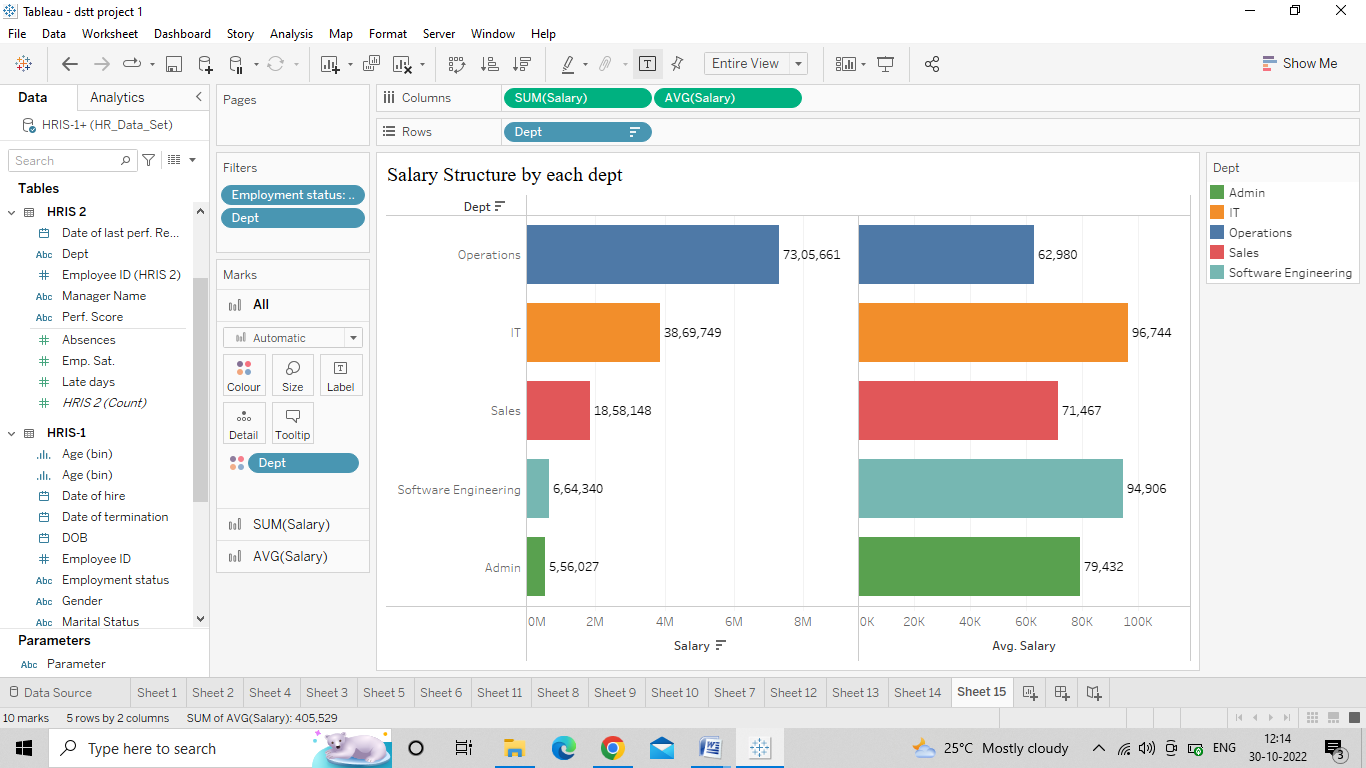
**2.Salary Structure**

1) What was the current total salary expense for each department?

● Bring Dept to rows and Salary to columns

● Duplicate the Salary field in the columns and convert the aggregation to ‘Avg’

● Filter for active employees and enable text labels



**Insight:**

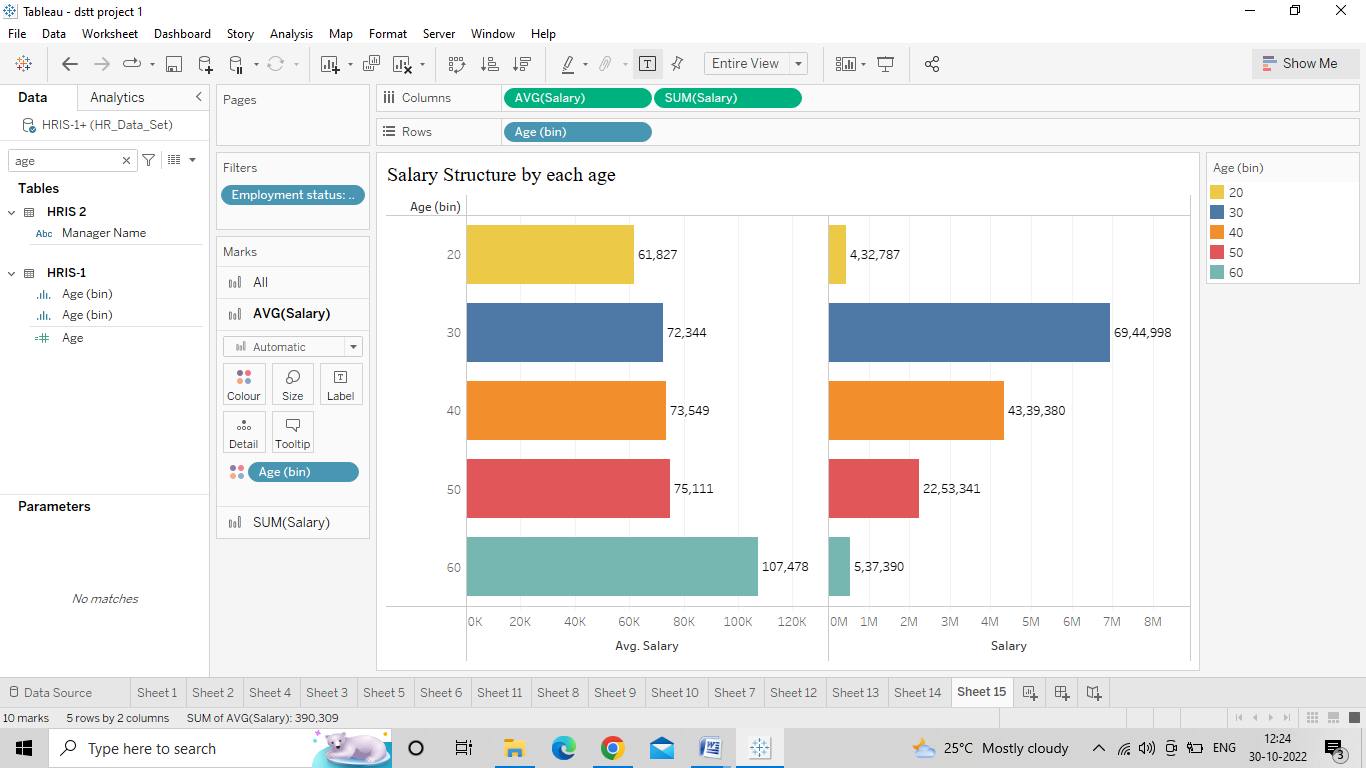
Among all departments excluding executive office, given it is the CEO, the department with average highest salary is IT and the department with the lowest average salary is Operations. However, in terms of the total salary expenses, Operations is on the top mainly because of the number of the employees in the department.

2) What is the salary structure by Age?

● Bring the Age (bin) to the columns and salary to the rows

● Duplicate salary in the rows and convert aggregation to Avg

● Filter for active employees and enable text labels



**Insight:**

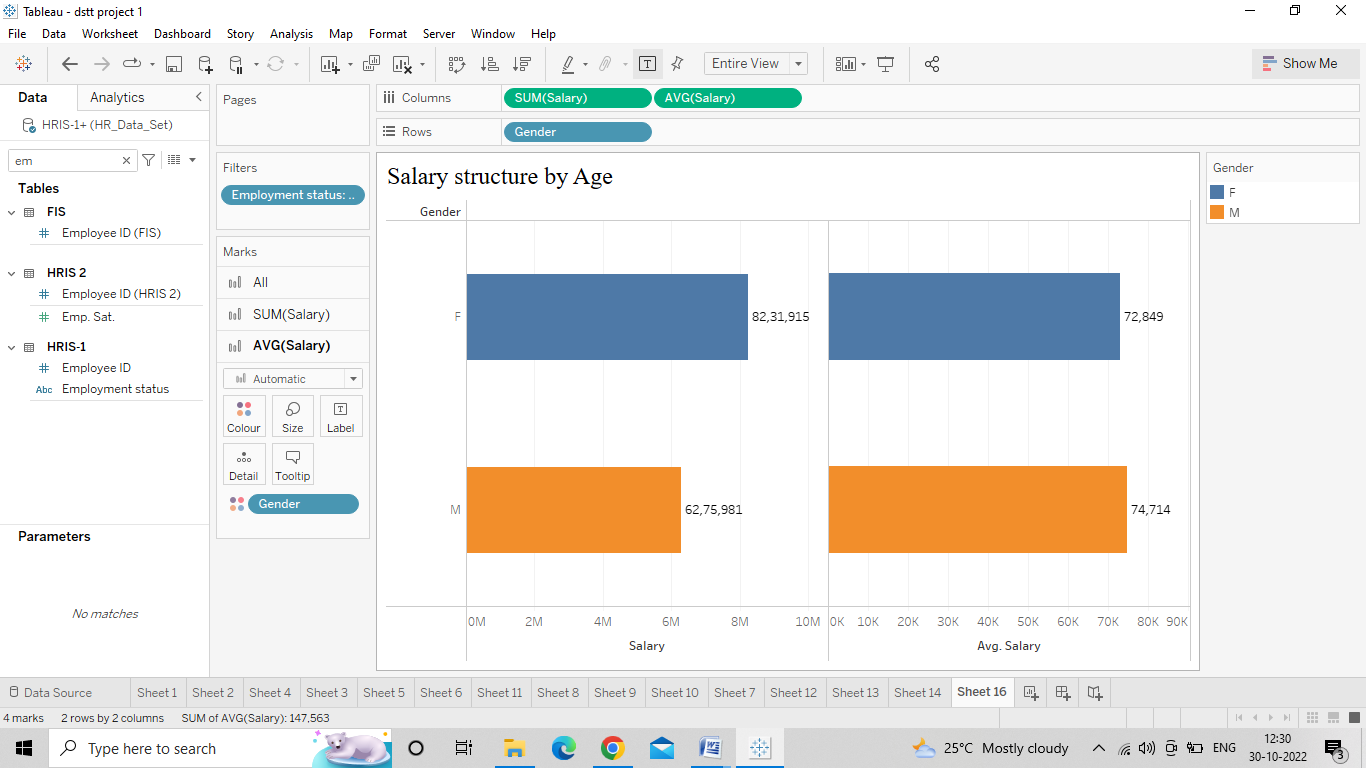
In terms of average salary, across all bins the salary levels were similar except the 20 – 30 year bucket and 60+. This is mostly because there are only few employees in the bucket. However, in terms of total expense, the highest expense is in the 30 – 40 year category mainly because of the number of employees in the group.

3) What is the salary structure by Gender?

● Bring Gender to the rows and Salary to the columns

● Duplicate Salary in the columns and convert the aggregation to show Avg

● Filter for active employees and show the text labels



**Insight:**

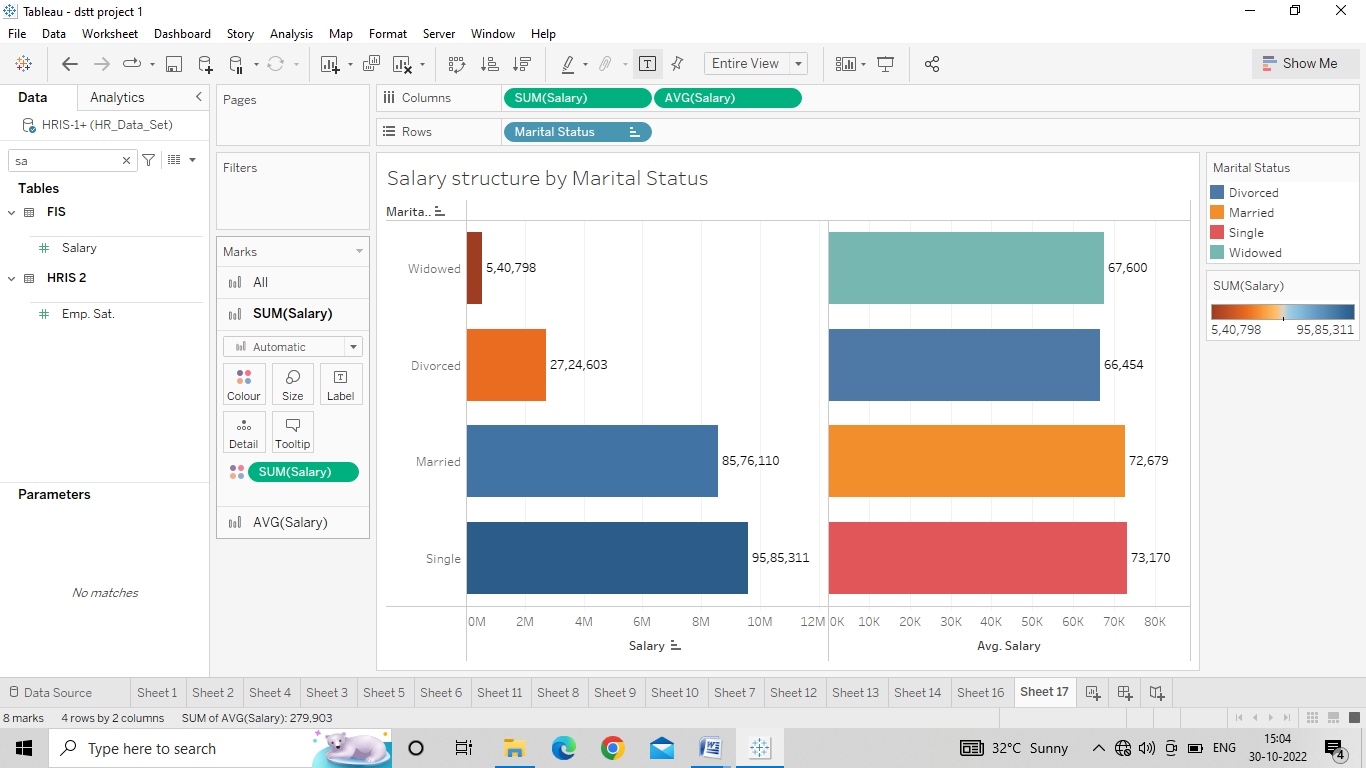
More salary expense goes to the female gender mostly due to the number of female staffs. In terms of the average salary, the difference in the salary made by female staff in comparison with that of the male staff is about < $2000.

4) What is the salary structure by Marital Status?

● Bring Marital Status to the rows and Salary to the columns

● Duplicate Salary in the rows and convert the aggregation to show Avg

● Filter for active employees and show the text labels



**Insights:**

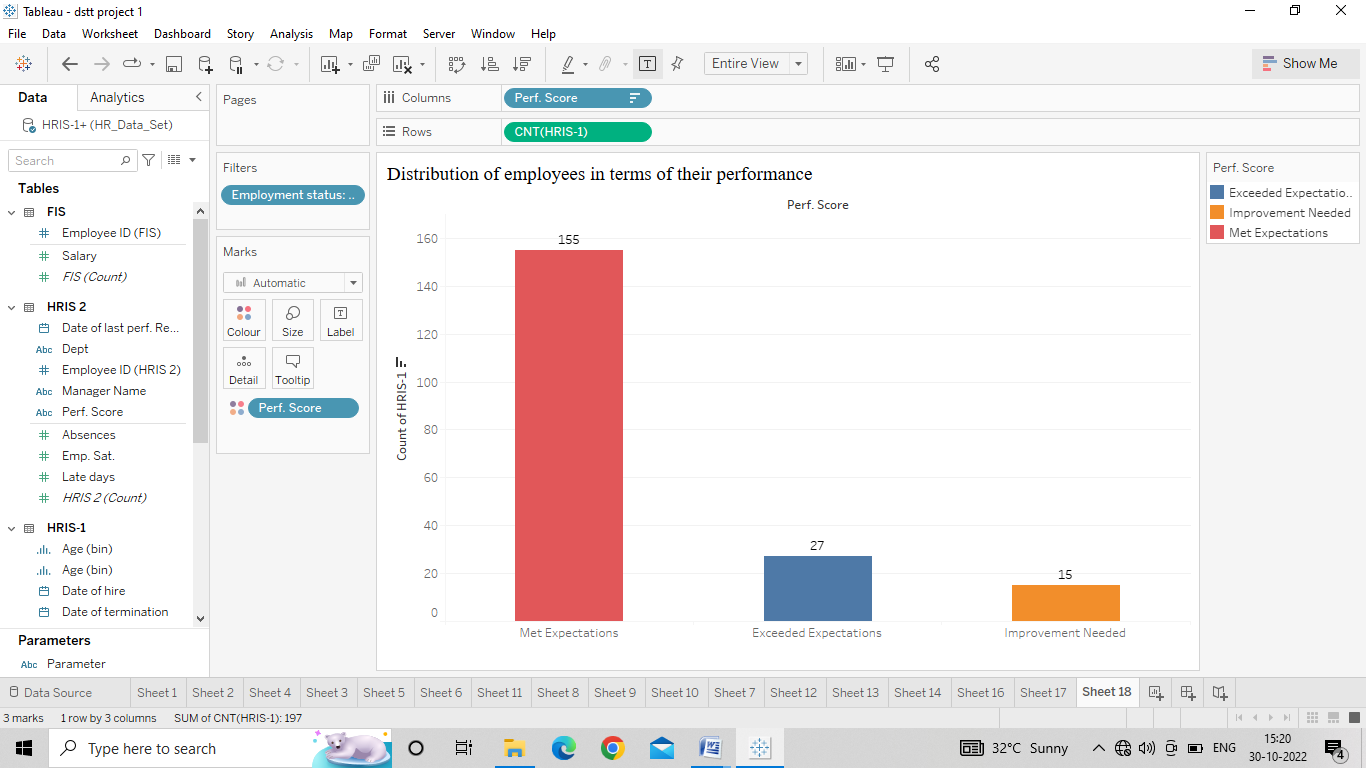
It seems like the average salaries made by single and married employees were similar however the divorced employees made ~$8000 lesser. This could possibly because this category has only few employees however given that there are 24 employees in this group, we cannot disregard the difference in average salary.

**Performance Results:**

1. What was the distribution of employees in terms of their performance?

● In order to create a distribution of employees by performance, bring ‘Perf. Scores’ to columns and HRIS (Count) to the rows

● Filter for active employees and show the text labels



*Insights:*

We see that most (155) employees met the expectations, while 27 employees exceeded the expectations and 15 employees did not meet the standards and had to improve their performance

1. Performance by Department

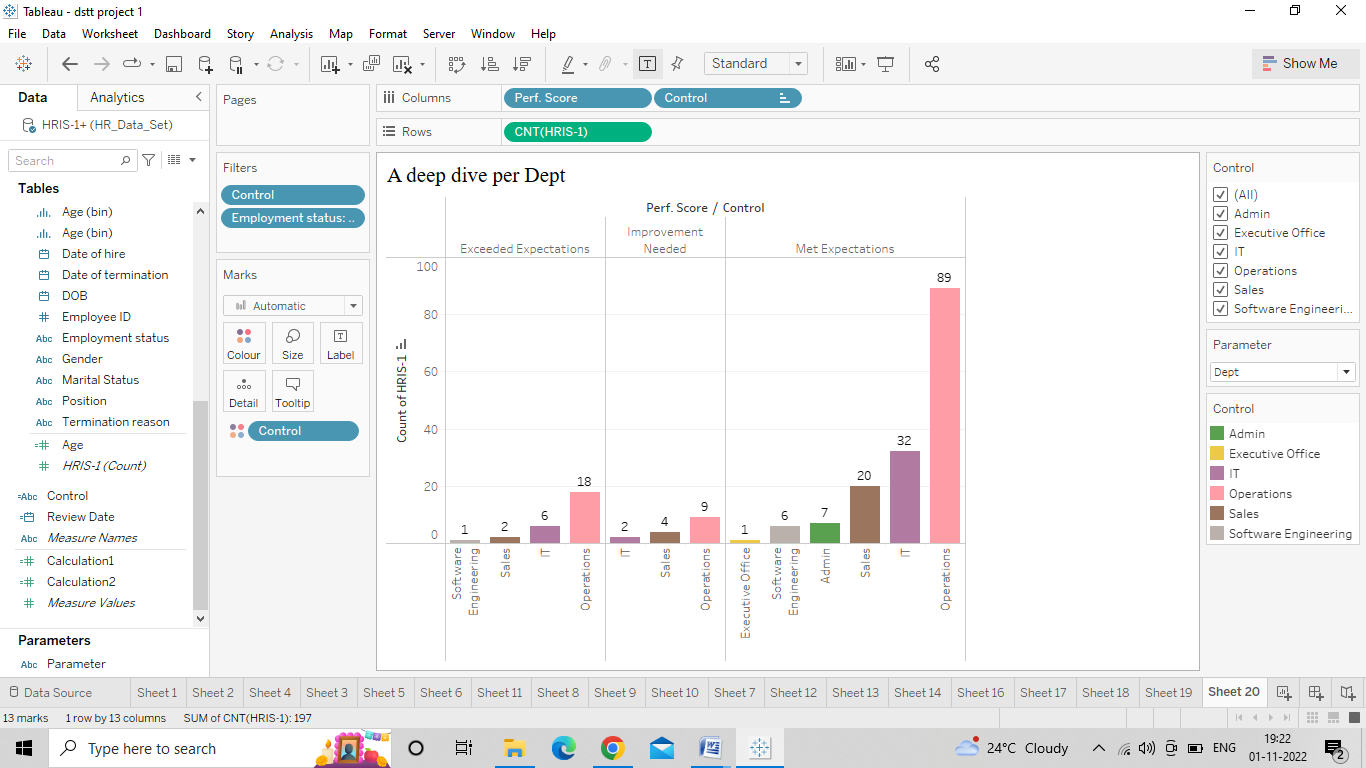
We will create a table using parameters to show the performance by different dimensions

● Bring ‘Perf Score’ to the columns and Dim Select Control to rows

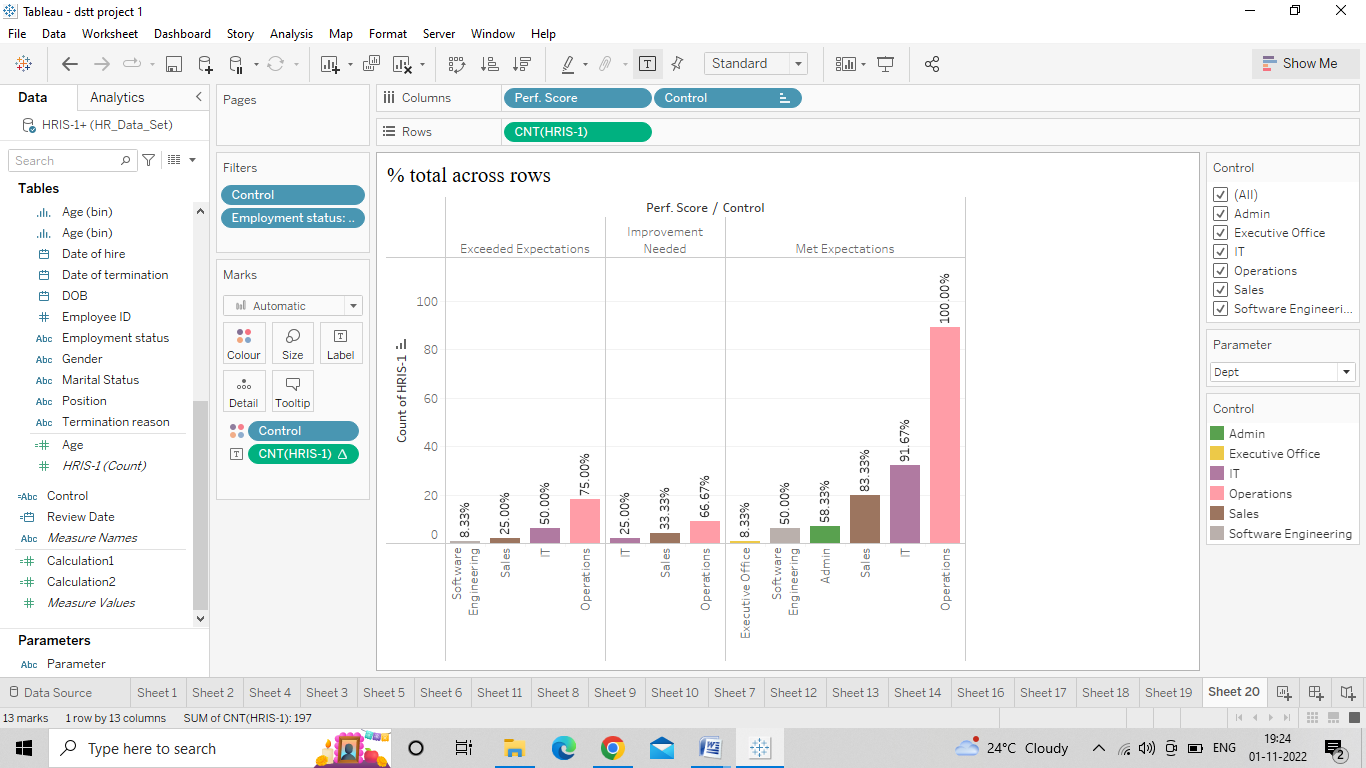
● Include the HRIS (count) in the text

● Select Dept from the parameter control

● Filter for active employees



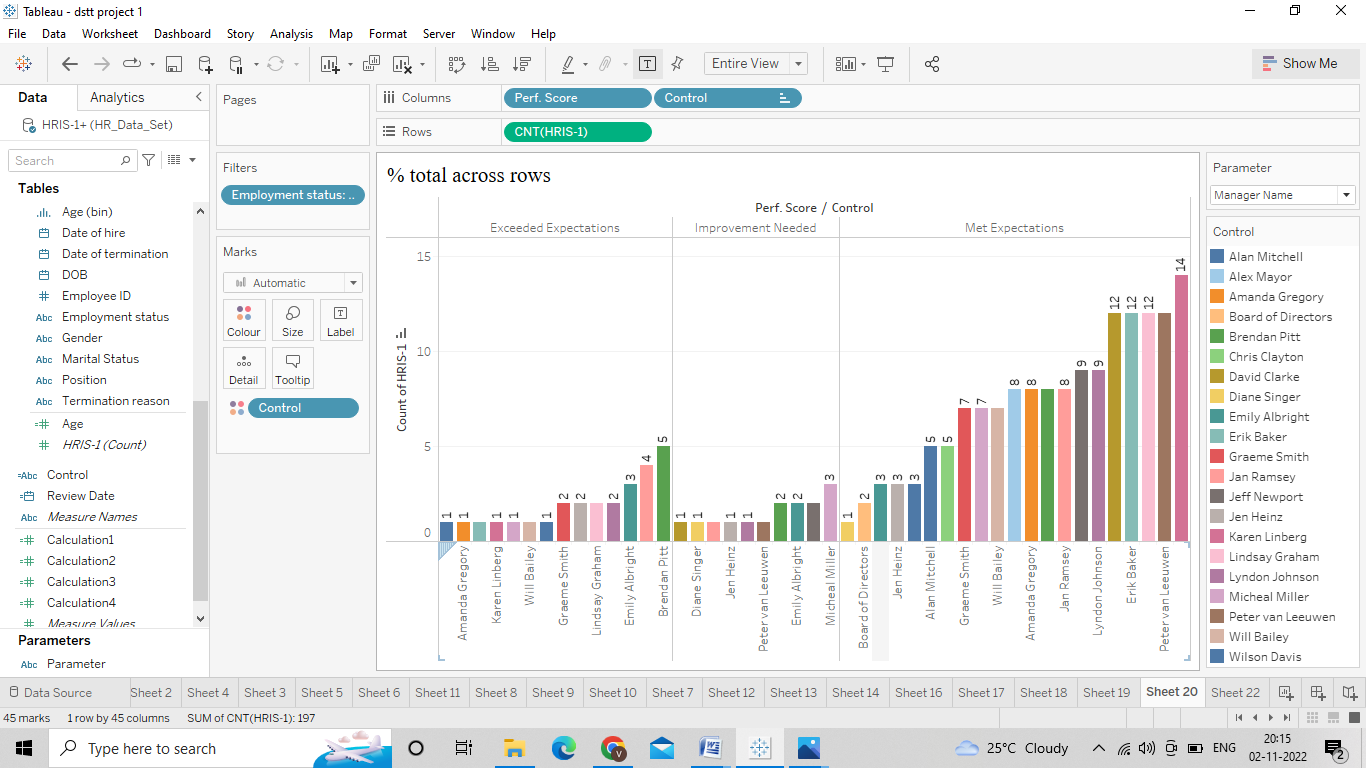
● Duplicate this and include table calculation in the text to show the % total across rows



1. Performance by Manager Name

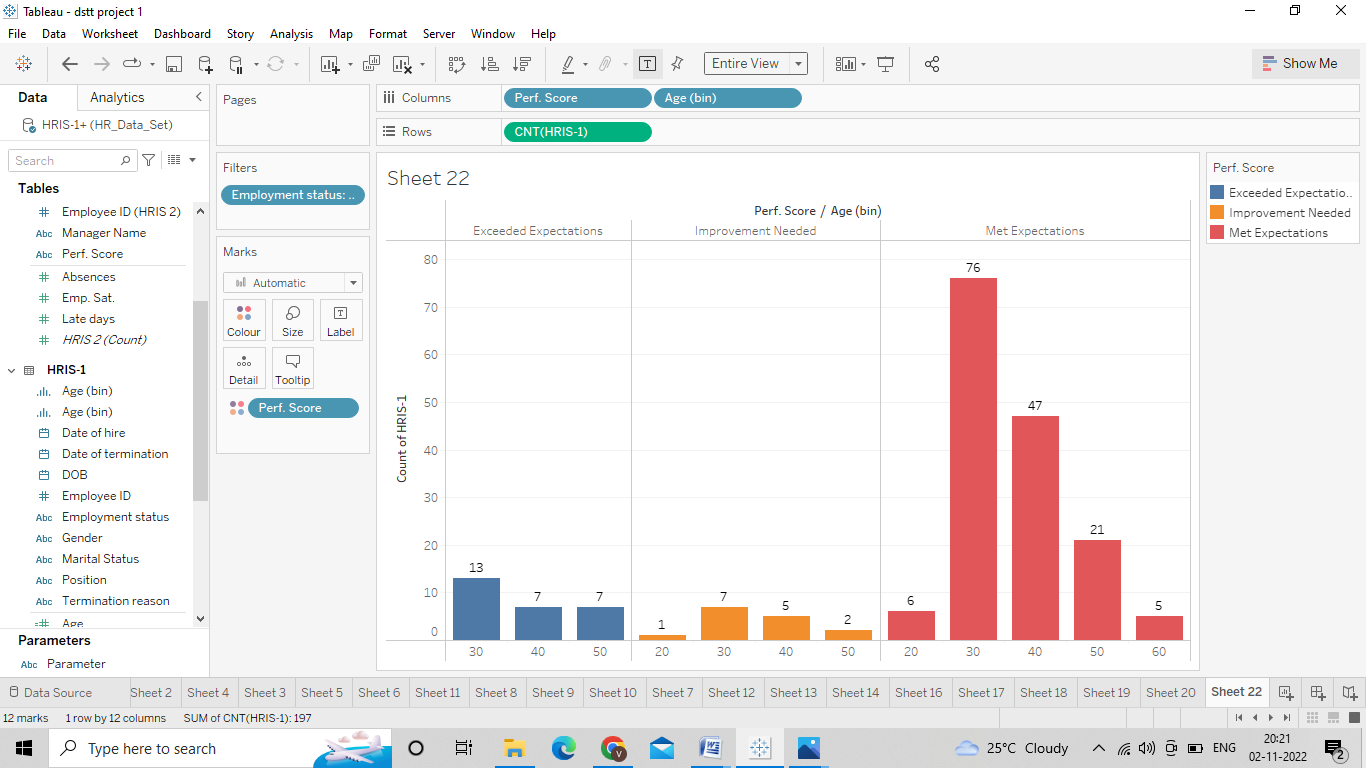
● Update the parameter and calculated field for parameter control to include the ‘Manager name’

● From the parameter control, select ‘Manager Name’



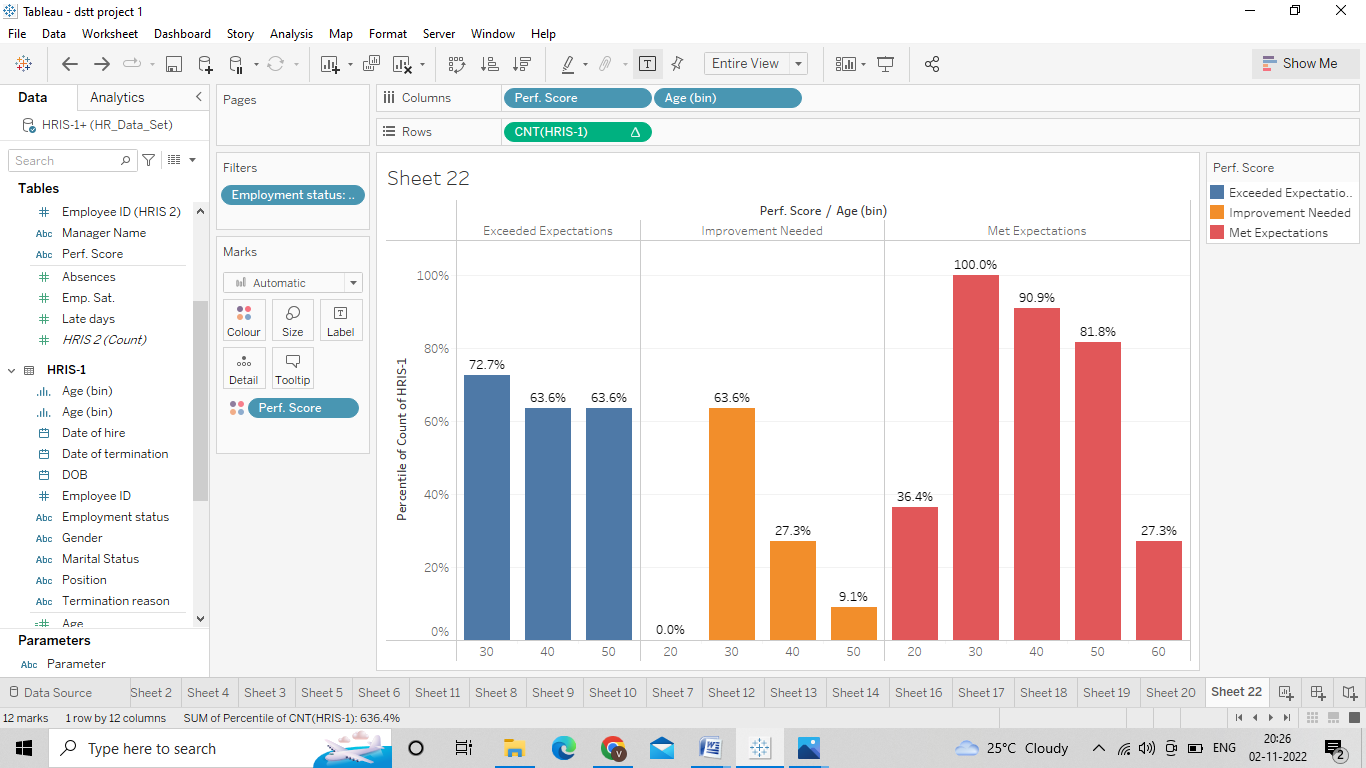
4.Performance by Age

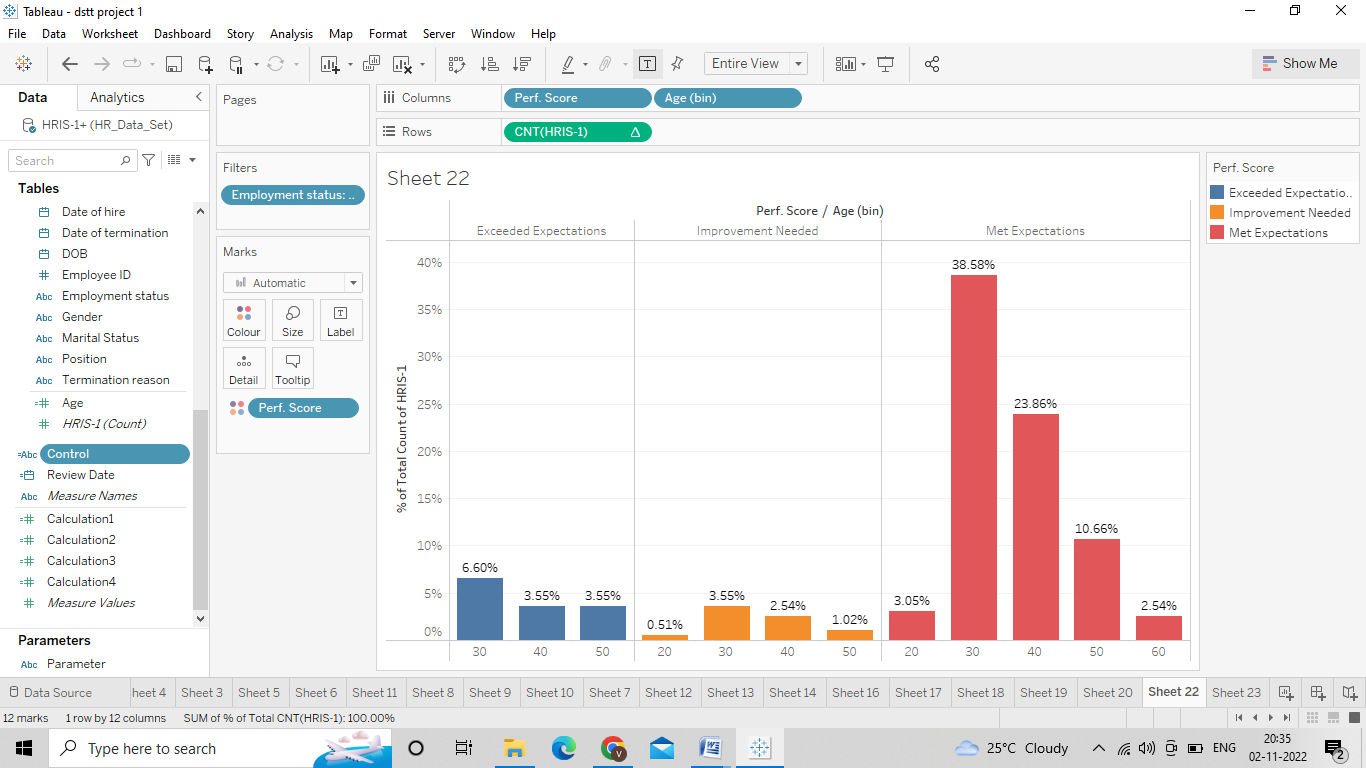
● In order to create a distribution of employees by performance and age, bring ‘Age (bins) to columns and HRIS (Count) to the rows, bring ‘Perf Scores’ to the color. We could also use the cross table created earlier.



● Include table calculation Change the HRIS (count) to show % total

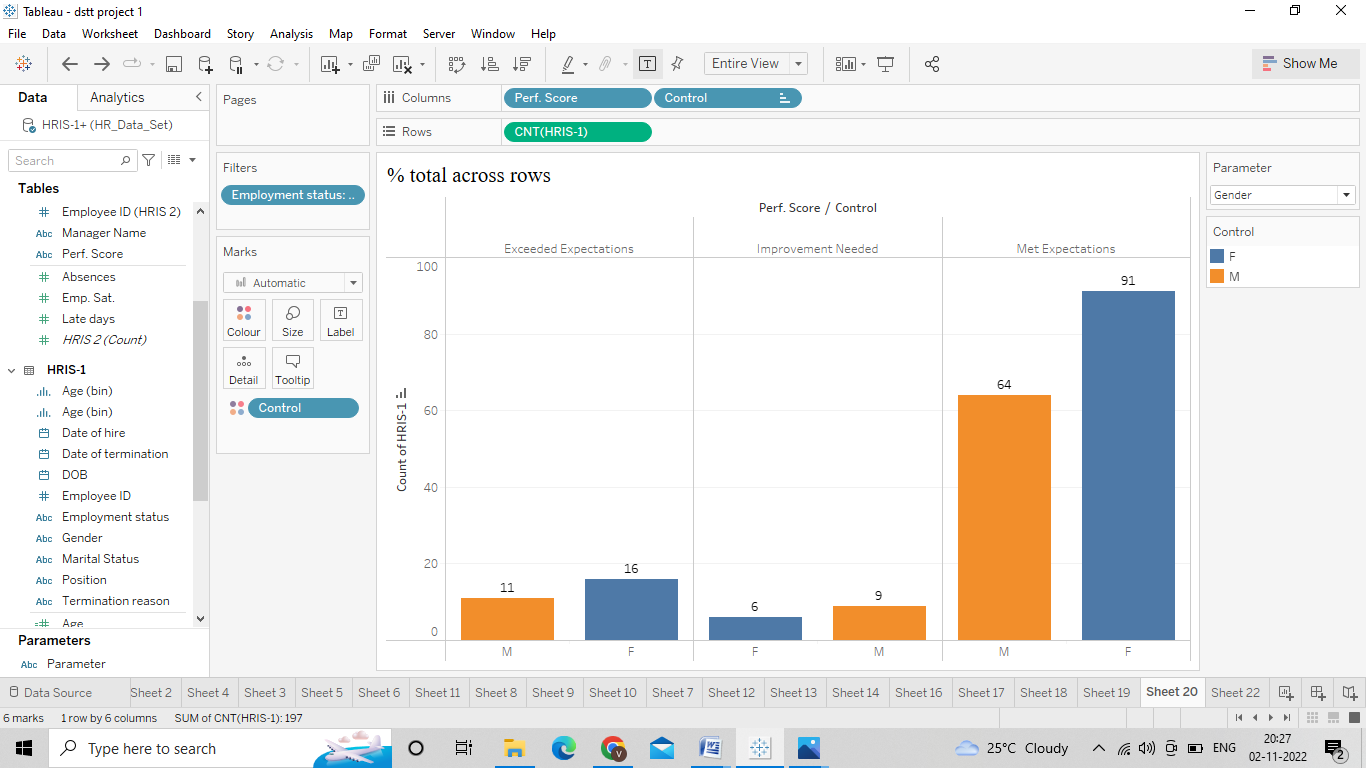
● Filter for active employees and show the text labels





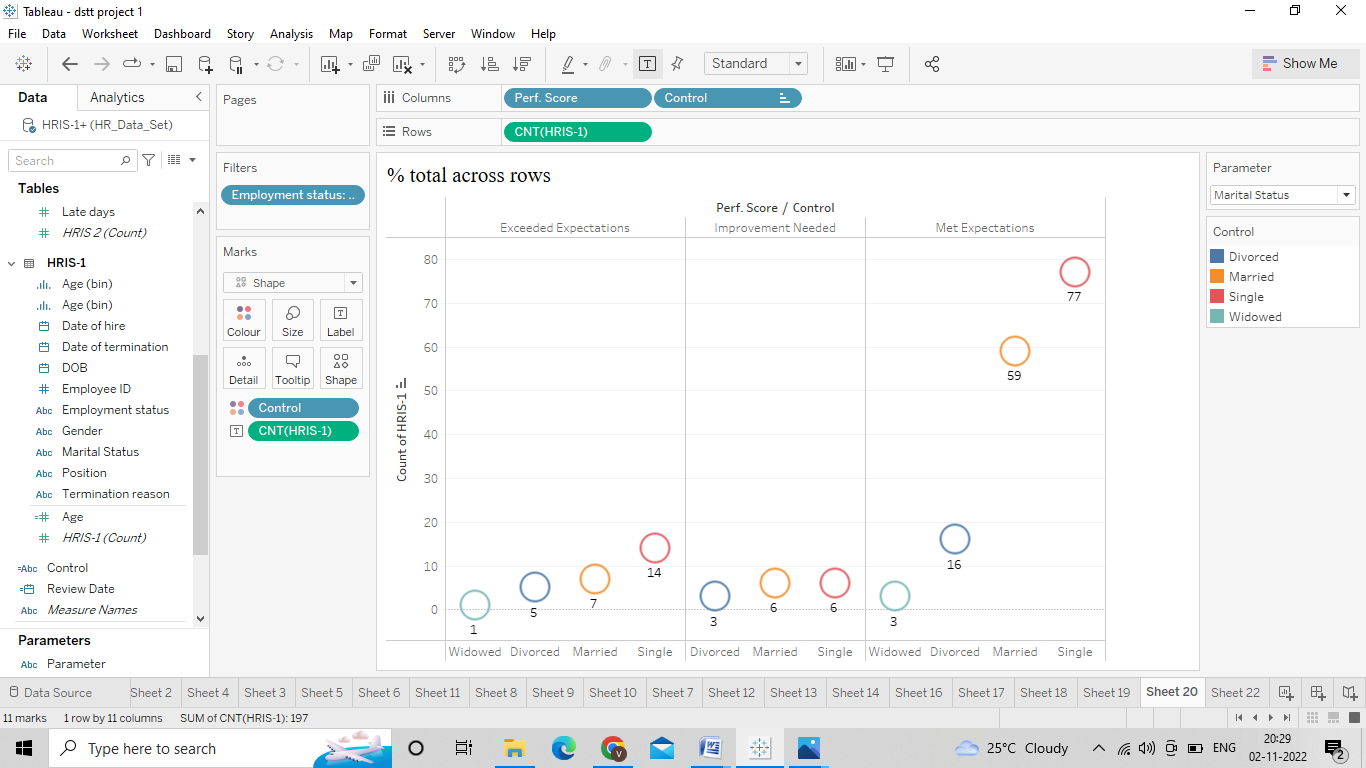
1. Performance by Gender

● From the dashboard created earlier, select ‘Gender’ from the parameter control



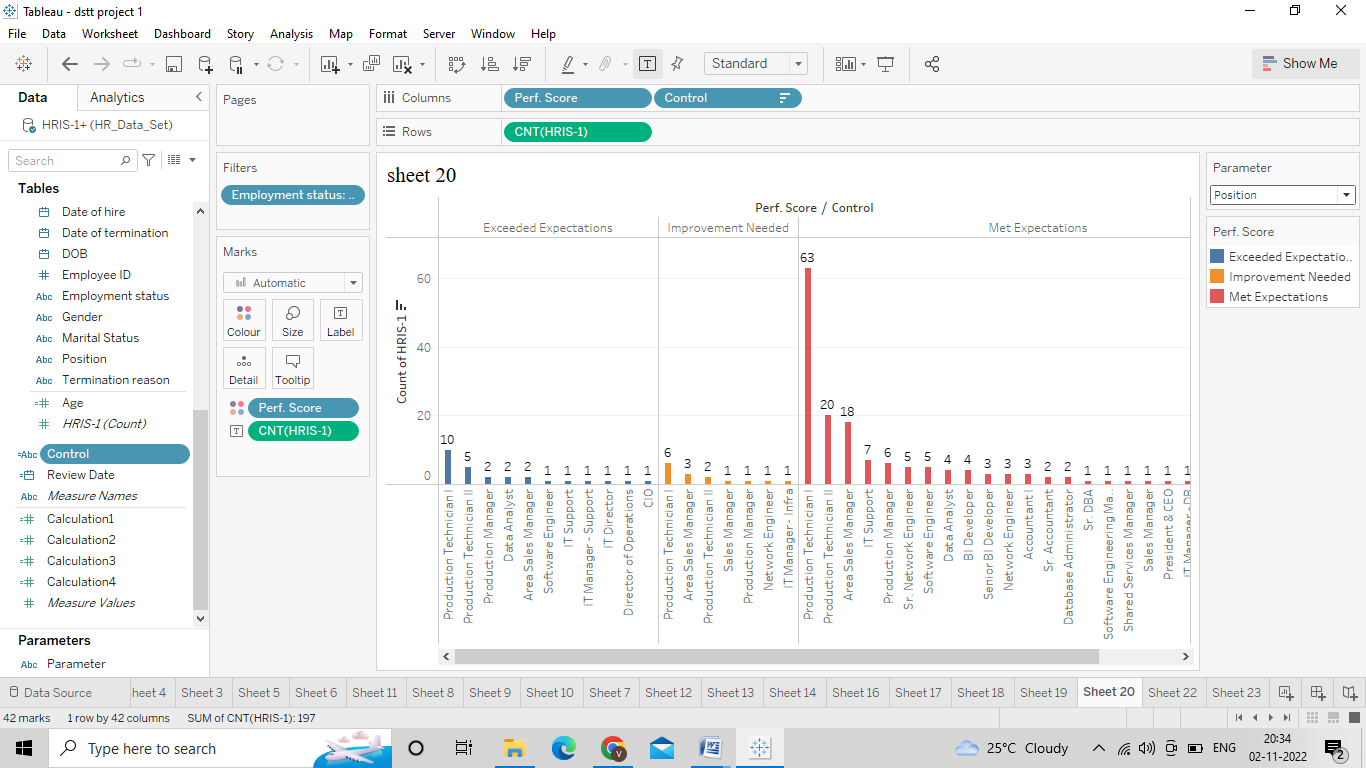
6) Performance by Marital Status

● From the dashboard created earlier, select ‘Marital Status from the parameter control



7) Performance by Position

● From the dashboard created earlier, select ‘Position’ from the parameter control

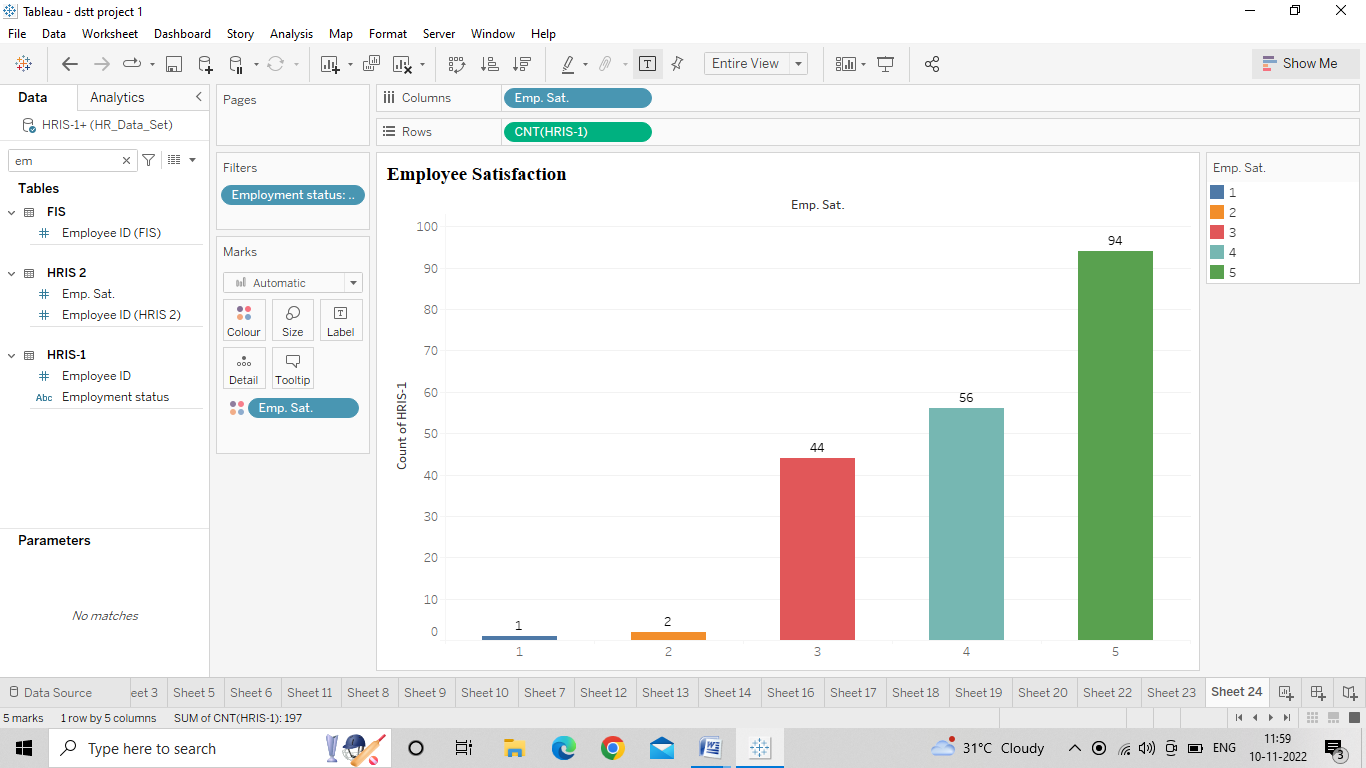


**Employee Satisfaction**

**How satisfied are the employees?**

● Create a chart similar to the ones earlier to show ‘Emp sat dim’ vs count of rows. Note that we convert the emp. Sat into dimension and used in this chart

● Filter for active employees and show text labels

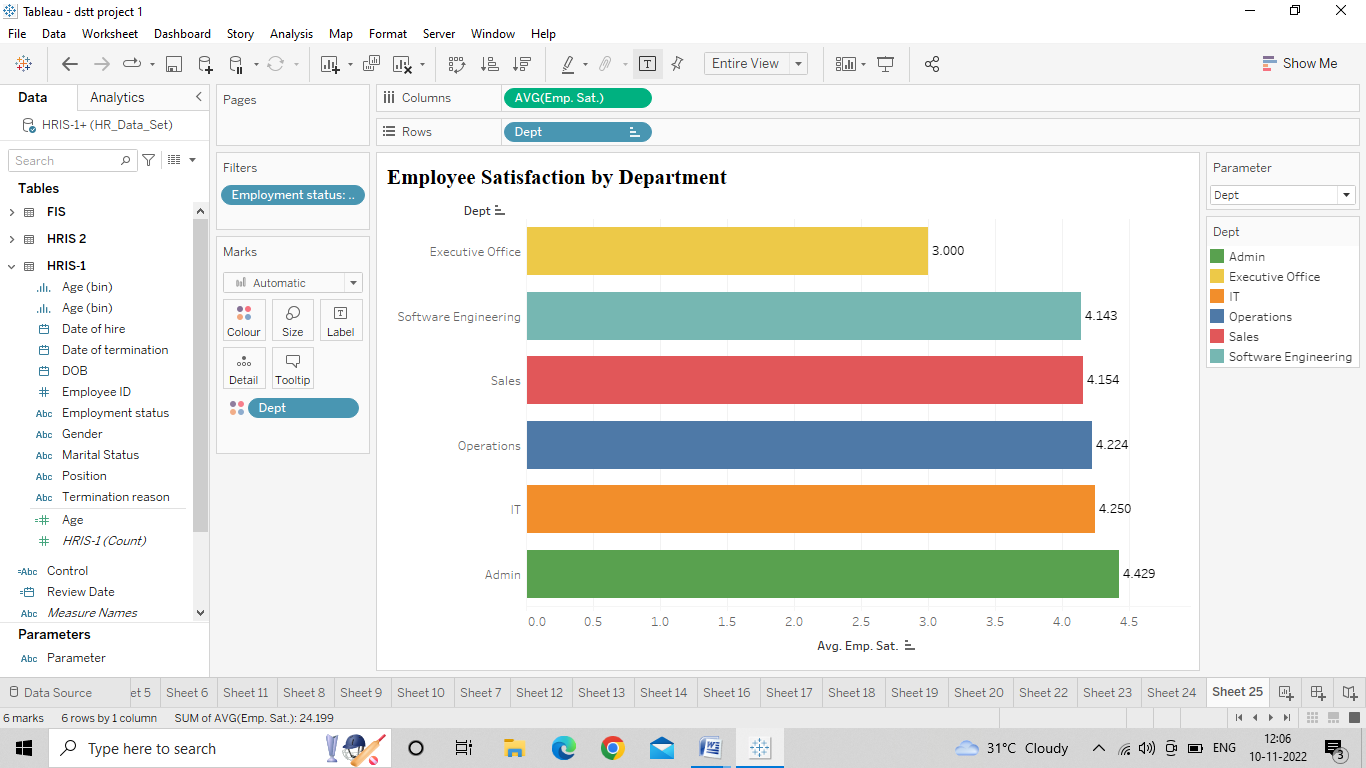


**Employee Satisfaction by Department**

● Similar to the previous chart, we create a chart showing average employee satisfaction Vs parameter control

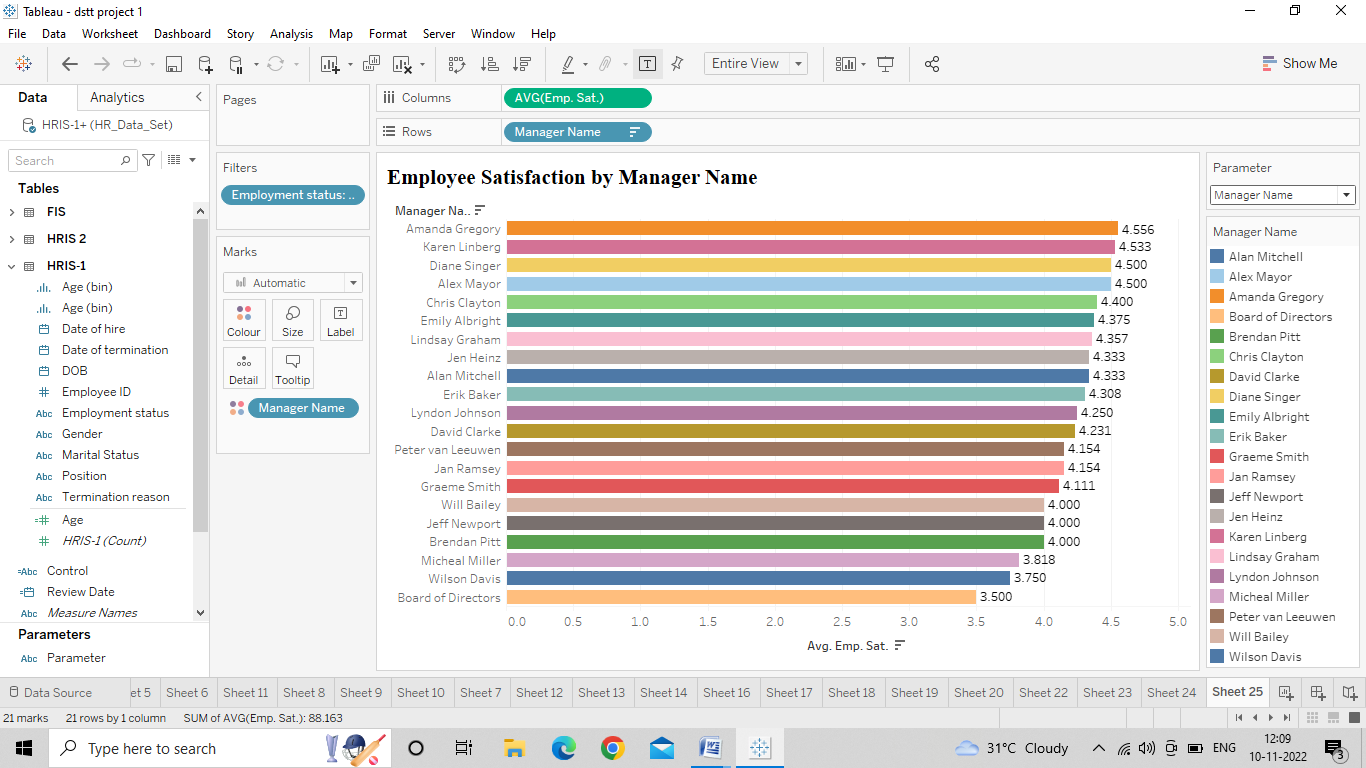
● Select ‘Dept’ from the parameter drop down

● Filter for active employees and show text labels

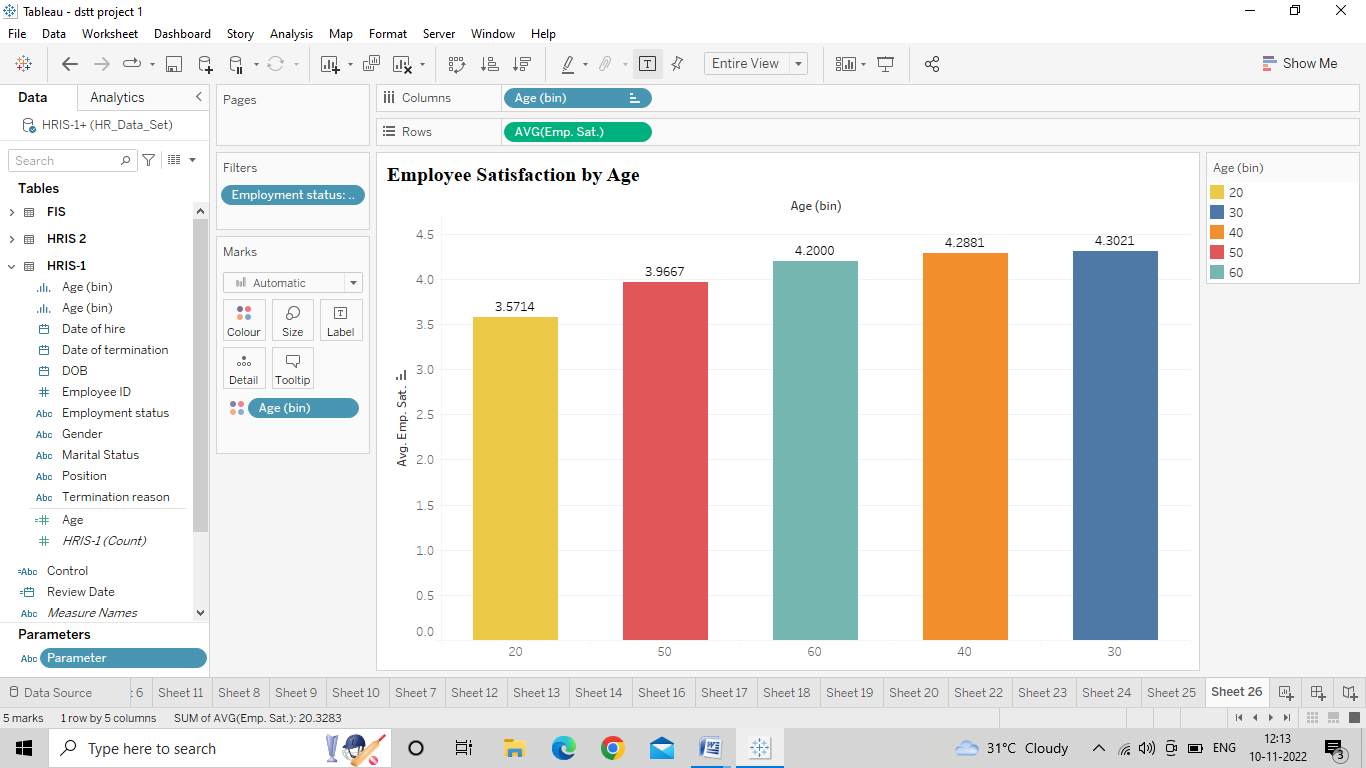
****

**Employee Satisfaction by Manager Name**

● Select ‘Manager name’ from the parameter drop down

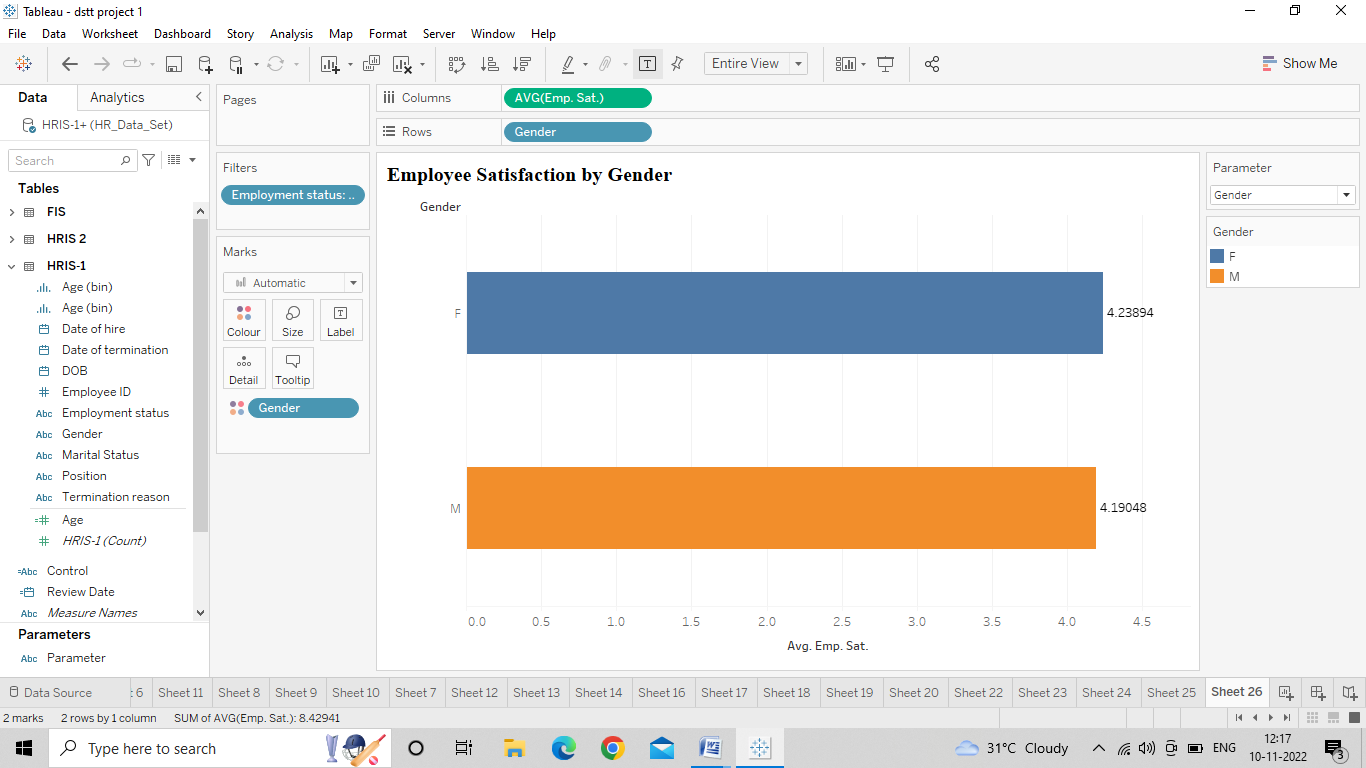


**Employee Satisfaction by Age**



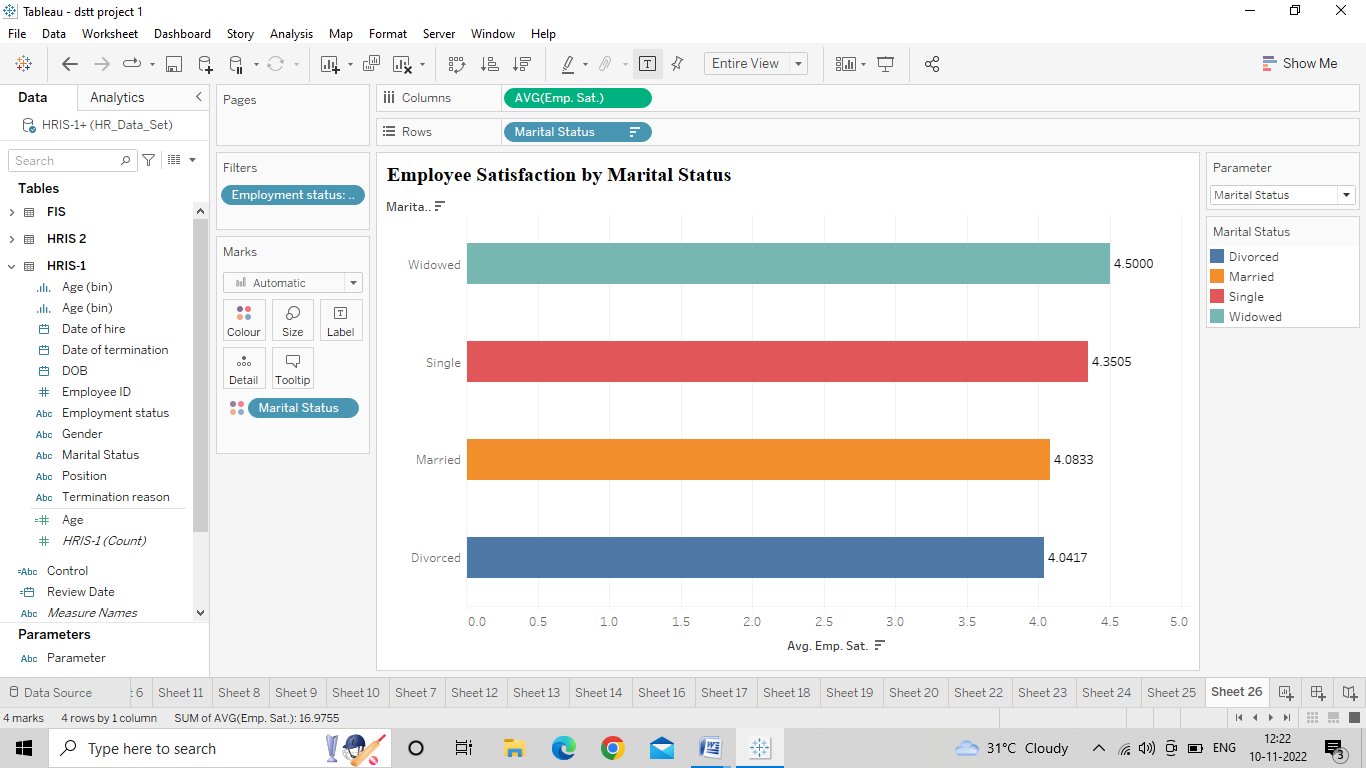
**Employee Satisfaction by Gender**

● Select ‘Gender’ from the parameter drop down



**Employee Satisfaction by Marital Status**

● Select ‘Marital Status’ from the parameter drop down

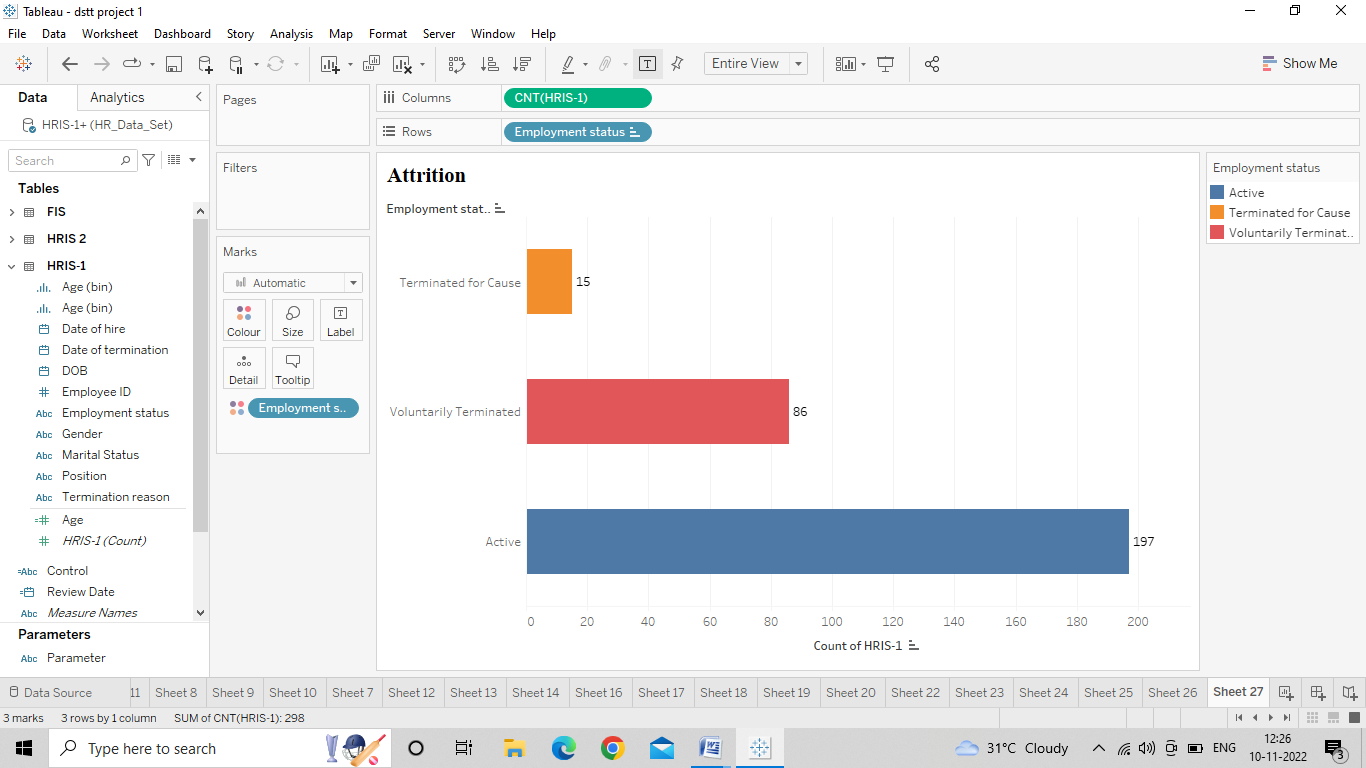


**5.Attrition**

How many employees have left the organization in total?

● Bring Employment status to rows and HRIS count to columns

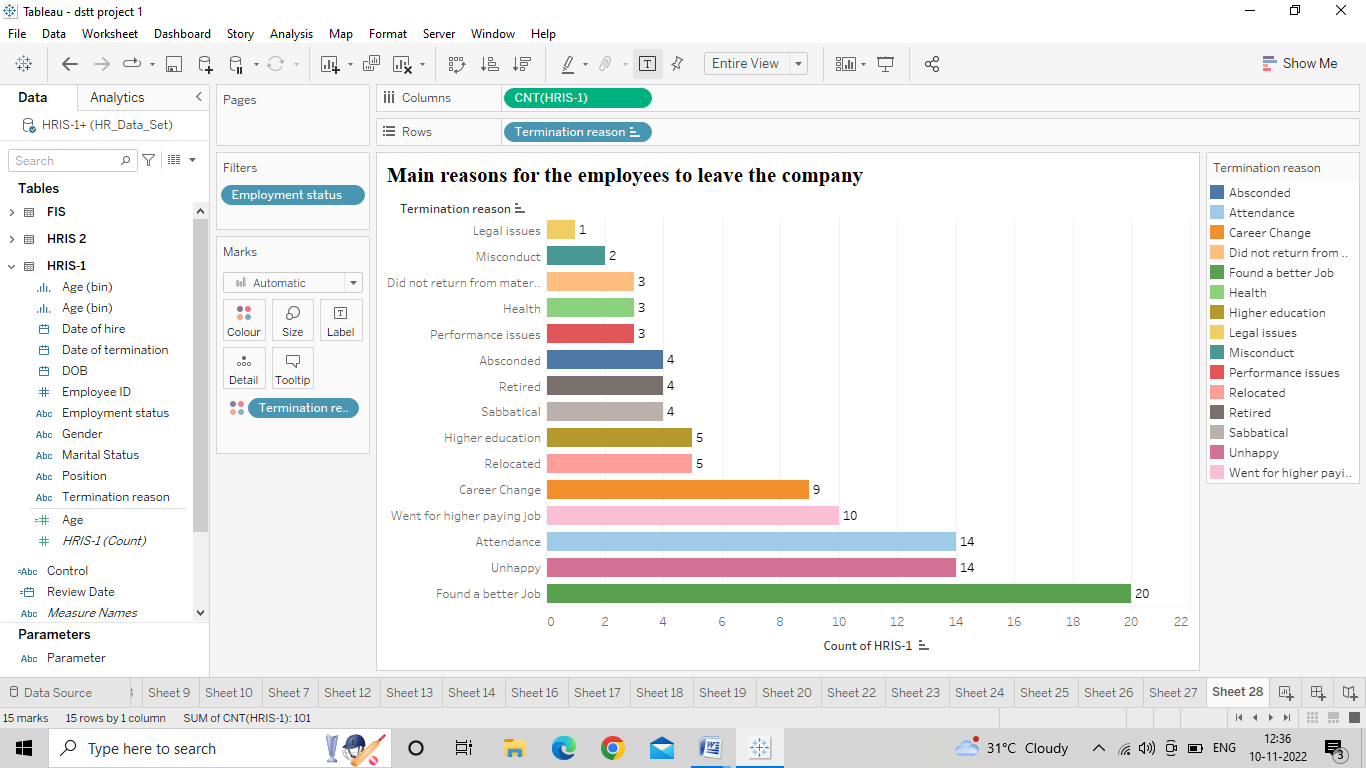
● Color the marks with employment status and show text labels



**What were the main reasons for the employees to leave the company?**

● Bring high-level and low-level termination reasons to the rows and HRIS count to the columns

● Filter the chart for terminated employees and enable text labels



**Voluntary Vs non-voluntary**

● Bring Employment Status to the color in the same chart

