

Keep Remove Rows in Power Query

Keep Remove row option is available under Home tab in Power Query. It provide various options to clean the data.

The screenshot displays the Power Query Editor window with the 'Home' tab selected. The 'Remove Rows' dropdown menu is open, showing options to clean data. A red rectangle highlights the 'Remove Rows' dropdown and its menu items. The data table has columns EID, NAME, and DOB. The 'Query Settings' pane on the right shows the 'APPLIED STEPS' list, which includes 'Removed Blank Rows' and 'Removed Duplicates'.

| EID | NAME | DOB |
|-----|-----------------------|------------|
| 1 | 1001 RAMESH GUPTA | 9/1/1990 |
| 2 | 1002 RAMESH GUPTA | 9/15/1995 |
| 3 | 1003 Rajesh Sharma | 3/16/2001 |
| 4 | 1004 Yogeshwar Sharma | xxx |
| 5 | 1005 Manoj Kumar | 7/1/1988 |
| 6 | 1006 Rohit Gupta | 3/31/1992 |
| 7 | 1007 Kapil Sharma | 9/28/1987 |
| 8 | 1008 Archana Sharma | 5/10/1985 |
| 9 | 1009 Ranjeeta Goyal | 12/31/1989 |
| 10 | 1010 Komal Singh | 3/31/1990 |
| 11 | 1011 AMIT KAPOOR | 1/1/1992 |
| 12 | 1012 MONIKA ARORA | 1/1/1982 |
| 13 | 1013 ROHAN KUMAR | 10/10/1998 |
| 14 | 1014 RAJAT KUMAR | 10/10/1998 |
| 15 | 1015 ABHISHEK SHARMA | 10/11/1982 |
| 16 | 1019 RAJAT KAPOOR | 10/11/1983 |
| 17 | 1020 Rohan Kumar | 10/10/1989 |
| 18 | 1021 Lalit Sharma | 10/12/1989 |
| 19 | 1022 Monika Bajaj | 10/12/1992 |
| 20 | 1023 Devender Kapoor | 10/10/1982 |
| 21 | 1024 Ravinder Kumar | 10/10/1985 |
| 22 | 1025 GAURAV GUPTA | 12/10/1987 |
| 23 | 1026 David | 12/17/1987 |
| 24 | 1027 KONIKA KAPOOR | 10/10/1980 |
| 25 | 1028 RAJAT SHARMA | 1/14/1989 |

3 COLUMNS, 63 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 7:38 AM

Type here to search

ENG 7:42 AM
IN 11/1/2020



Manipulating Columns in Power Query

The Manage Column option under the Home tab allows us to Choose the column, Remove Column, Navigate to a particular column.

The screenshot shows the Microsoft Power Query Editor interface. The 'Home' tab is selected, and the 'Manage Columns' button in the ribbon is highlighted with a red rectangle. The ribbon also shows 'Choose Columns', 'Remove Columns', 'Keep Rows', 'Remove Rows', 'Sort', 'Split Column', 'Group By', 'Data Type', 'Merge Queries', 'Append Queries', 'Combine Files', 'Text Analytics', 'Vision', and 'Azure Machine Learning'. The main area displays a table with columns: EID, NAME, FNAME, LNAME, and ADDRESS. The table contains 27 rows of employee data. On the right, the 'Query Settings' pane is open, showing the 'Properties' tab with the query name 'emp' and a list of 'Applied Steps' including 'Source', 'Navigation', 'Promoted Headers', 'Changed Type', 'Inserted Age', 'Inserted Total Years', 'Changed Type1', 'Added Conditional Column', 'Added Conditional Column1', and 'Removed Columns'.

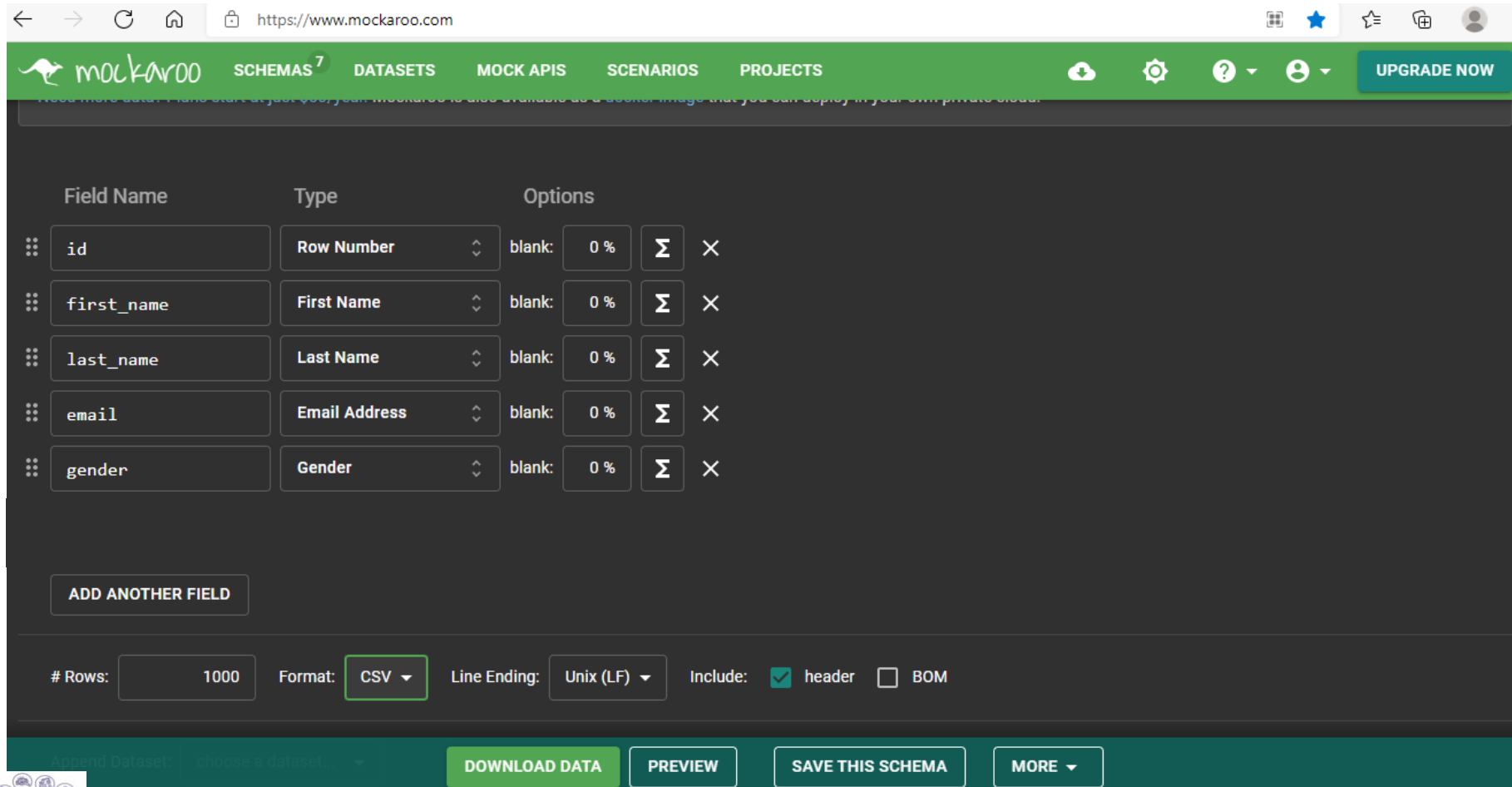
| EID | NAME | FNAME | LNAME | ADDRESS | |
|-----|------|------------------|-----------|---------|--|
| 1 | 1001 | RAMESH GUPTA | RAMESH | GUPTA | SECTOR 7,Rohini,Gurgaon |
| 2 | 1002 | Sandeep Sharma | Sandeep | Sharma | SECTOR 2,PALAM,GURGAON |
| 3 | 1003 | Rajesh Sharma | Rajesh | Sharma | Sector 11,Dwarka,Delhi |
| 4 | 1004 | Yogeshwar Sharma | Yogeshwar | Sharma | Sector 11,Dwarka,Delhi |
| 5 | 1005 | Manoj Kumar | Manoj | Kumar | B320,Janakpuri,Delhi |
| 6 | 1006 | Rohit Gupta | Rohit | Gupta | A32/620 Main Road,Sector 56,Noida |
| 7 | 1007 | Kapil Sharma | Kapil | Sharma | Jay Maa Apartments Flat 32,Sector 72,Gurgaon |
| 8 | 1008 | Archna Sharma | Archna | Sharma | 5C Karanpur,NP Road,Dehradun |
| 9 | 1009 | Ranjeeta Goyal | Ranjeeta | Goyal | Flat No 6A,Rajpur Road,Dehradun |
| 10 | 1010 | Komal Singh | Komal | Singh | Sector 64,Old Gurgaon Road,Gurgaon |
| 11 | 1011 | AMIT KAPOOR | AMIT | KAPOOR | B2/45 DDA FLATS,JANAKPURI,DELHI |
| 12 | 1012 | MONIKA ARORA | MONIKA | ARORA | B1/45,JANAKPURI,DELHI |
| 13 | 1013 | ROHAN KUMAR | ROHAN | KUMAR | B305,RAJOURI GARDEN,DELHI |
| 14 | 1014 | RAJAT KUMAR | RAJAT | KUMAR | HNO-B205,SECTOR 62,NOIDA |
| 15 | 1015 | ABHISHEK SHARMA | ABHISHEK | SHARMA | C2-101,JANAKPURI,DELHI |
| 16 | 1019 | RAJAT KAPOOR | RAJAT | KAPOOR | B501/1 POCKET 1,JANAKPURI,DELHI |
| 17 | 1020 | Rohan Kumar | Rohan | Kumar | B-256,RAJOURIGARDEN,DELHI |
| 18 | 1021 | Lalit Sharma | Lalit | Sharma | C2-401,Janak Puri,DELHI |
| 19 | 1022 | Monika Bajaj | Monika | Bajaj | DLF,Phase 5,GURGAON |
| 20 | 1023 | Devender Kapoor | Devender | Kapoor | E102/3 Pocket 1 Sector 3,Dwarka,Delhi |
| 21 | 1024 | Ravinder Kumar | Ravinder | Kumar | B302-Jay Maa Apt,Dwarka,Delhi |
| 22 | 1025 | GAURAV GUPTA | GAURAV | GUPTA | D41-Jay Maa Apt,Dwarka,Delhi |
| 23 | 1026 | David | null | David | M121/4,Vatika Tower,Gurgaon |
| 24 | 1027 | KONIKA KAPOOR | KONIKA | KAPOOR | C2-431,SUPERTECH,NOIDA |



Mackaroo- Random Data Generator

It is a free data generator tool which lets you create custom datasets as per the specified schema to test and demo your application.

<https://www.mockaroo.com>



The screenshot displays the Mockaroo web application interface. At the top, there is a navigation bar with links for SCHEMAS, DATASETS, MOCK APIS, SCENARIOS, and PROJECTS. Below this, a table defines the schema for a dataset with five fields: id, first_name, last_name, email, and gender. Each field has a specific type and options for blank values and a summary icon. At the bottom, there are controls for the number of rows (set to 1000), the output format (set to CSV), line ending (set to Unix (LF)), and whether to include a header (checked) or BOM (unchecked). Buttons for DOWNLOAD DATA, PREVIEW, SAVE THIS SCHEMA, and MORE are visible at the bottom right.

| Field Name | Type | Options |
|------------|---------------|------------------------------|
| id | Row Number | blank: 0 % Σ \times |
| first_name | First Name | blank: 0 % Σ \times |
| last_name | Last Name | blank: 0 % Σ \times |
| email | Email Address | blank: 0 % Σ \times |
| gender | Gender | blank: 0 % Σ \times |

ADD ANOTHER FIELD

Rows: 1000 Format: CSV Line Ending: Unix (LF) Include: ☒ header ☐ BOM

Append Dataset: choose a dataset

DOWNLOAD DATA PREVIEW SAVE THIS SCHEMA MORE



Mackaroo- Random Data Generator

Download dataset using Mackroo.com to create the following tables:

Student

Student ID

Student Name

Gender

DOB

Birth Country

Citizenship

Status (**Active**/Deferred/Withdrawn)

Reason for withdrawal
(Homesick /Academically
Unprepared/Expensive)

Feedback with 5 questions

SID

Q01 - Professors teach well at this university?

Q02 - Teacher encourage us to perform better?

Q03 - I feel job-competent moving forward
from this university?

Q04- Lecturers were proficient enough?

Q05 - I would like to recommend this
educational institution to others?

Responses : Strongly Agree - 5

Moderately Agree - 4

Neither Agree or Disagree - 3

Moderately Disagree - 2

Strongly Disagree - 1



Data Modelling in Power BI

Data Modeling is used to connect multiple data sources using relationship and allows us to create visuals based on multiple data sources.

To create a data model we need to switch to Model view.



Report View - Allows us to create the visuals



Data View – to see the data from various sources

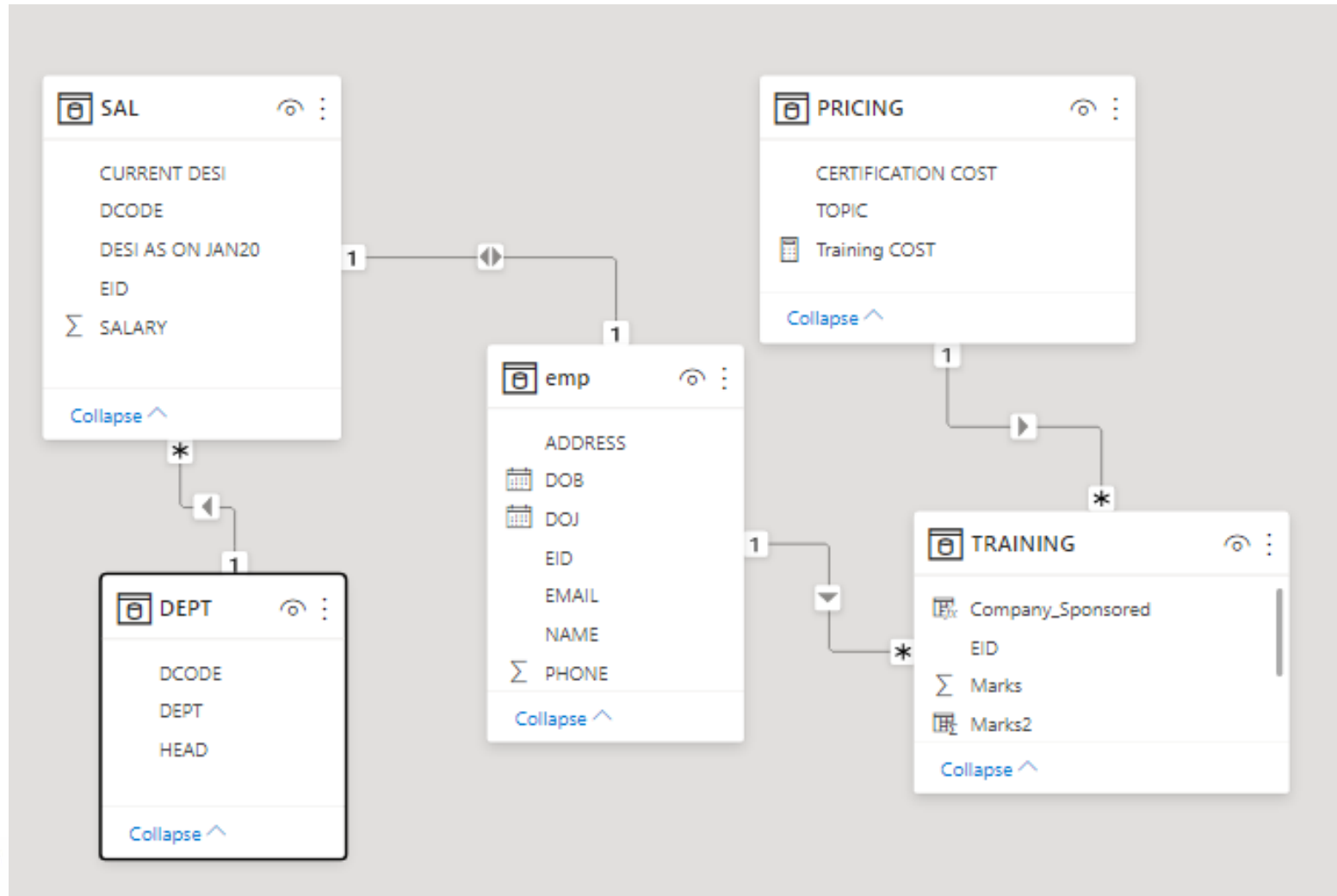


Model View – to manage relationships



Data Modelling in Power BI

The below model shows the relation ship between Employee, Salary, Dept, Pricing & Training Tables.



Data Modelling in Power BI

The objective of the model is to analyze the training status on the basis of department head

Training Analysis:

- No of Training applied under each head

- No of training completed under each head

- Training result of each head

- No. of company Sponsored Trainings under each head (*Marks ≥ 95 is CS*)

- Training applied for each module

- Percentage of self & company sponsored training

Cost Analysis:

- Department wise total salary

- Department wise cost of company sponsored training

- Bifurcation of company & self sponsored training under each module

- Total no of trainings conducted under each module along with the total cost.



Data Modelling in Power BI



Data View – to see the data from various sources, new columns & measures can also be created in this view.

> Column Tools – New Column

The screenshot displays the Power BI Data View interface. The 'Column tools' ribbon is active, showing options for Name, Format, Summarization, Data category, Sort by column, Data groups, Manage relationships, and New column. The 'New column' button is highlighted. Below the ribbon, a formula bar contains the DAX formula: `1 RESULT = IF(TRAINING[FINALMARKS]>=85, "CLEAR", IF(TRAINING[FINALMARKS]=0, "NA", "RE-APPEAR"))`. The main area shows a table with columns: EID, MODULE, STATUS, Marks, RESULT, and FINALMARKS. The 'RESULT' column is highlighted in yellow. The 'Fields' pane on the right shows a search bar and a list of fields: DEPT, DCODE, DEPT, HEAD, emp, PRICING, SAL, TRAINING, EID, FINALMARKS, Marks, MODULE, RESULT, and STATUS. The 'RESULT' field is selected.

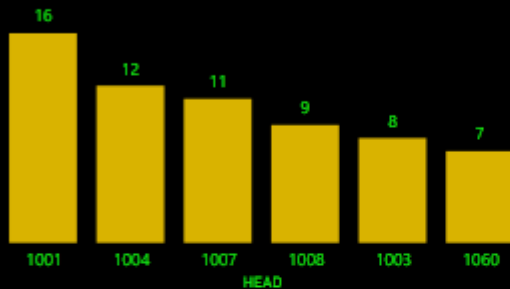
| EID | MODULE | STATUS | Marks | RESULT | FINALMARKS |
|-------|----------|---------|-------|--------|------------|
| E1065 | TABLEAU | Pending | 81 | NA | 0 |
| E1053 | TABLEAU | Pending | 66 | NA | 0 |
| E1052 | TABLEAU | Pending | 82 | NA | 0 |
| E1038 | SQL | Pending | 83 | NA | 0 |
| E1003 | POWER BI | Pending | 70 | NA | 0 |
| E1003 | SQL | Pending | 79 | NA | 0 |
| E1022 | TABLEAU | Pending | 89 | NA | 0 |
| E1029 | SQL | Pending | 97 | NA | 0 |
| E1040 | POWER BI | Pending | 62 | NA | 0 |
| E1041 | POWER BI | Pending | 75 | NA | 0 |
| E1056 | SAS | Pending | 93 | NA | 0 |
| E1034 | TABLEAU | Pending | 62 | NA | 0 |
| E1056 | SAS | Pending | 68 | NA | 0 |
| E1028 | TABLEAU | Pending | 62 | NA | 0 |
| E1035 | POWER BI | Pending | 95 | NA | 0 |
| E1066 | POWER BI | Pending | 68 | NA | 0 |
| E1035 | TABLEAU | Pending | 62 | NA | 0 |
| E1044 | POWER BI | Pending | 61 | NA | 0 |
| E1058 | POWER BI | Pending | 93 | NA | 0 |



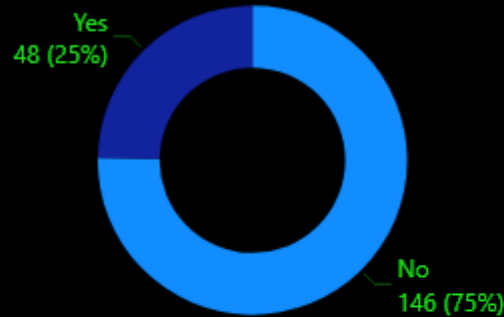
Data Modelling in Power BI

Training Status

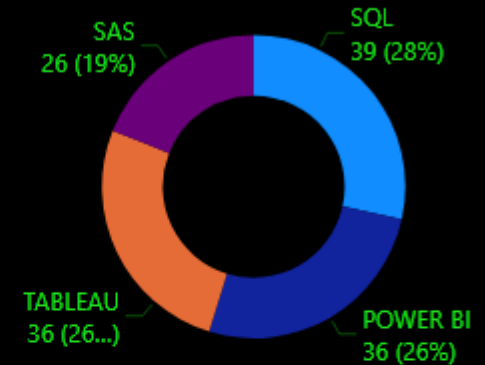
No. of Employee by HEAD



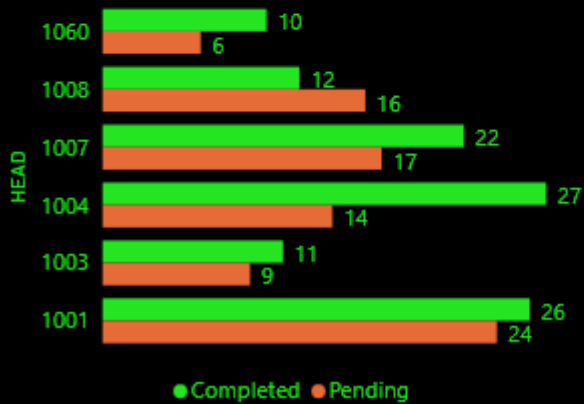
Count of EID by Company_Sponsored



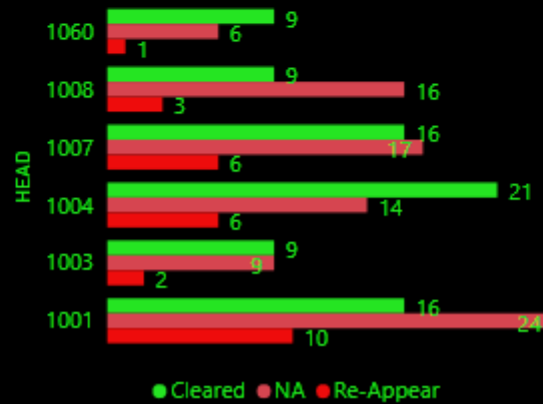
Count of EID by MODULE



Training Completion Status



Training Result



Training Cleared

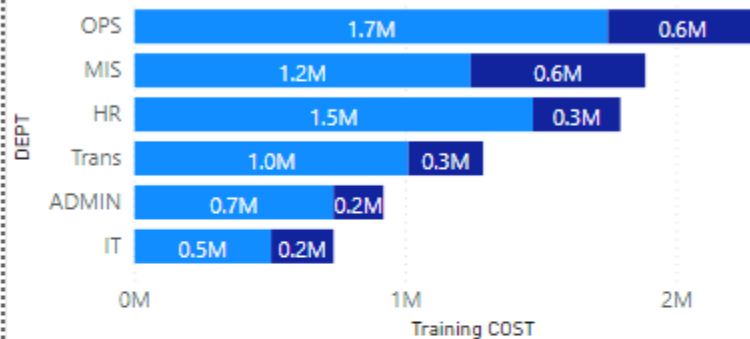


Data Modelling in Power BI

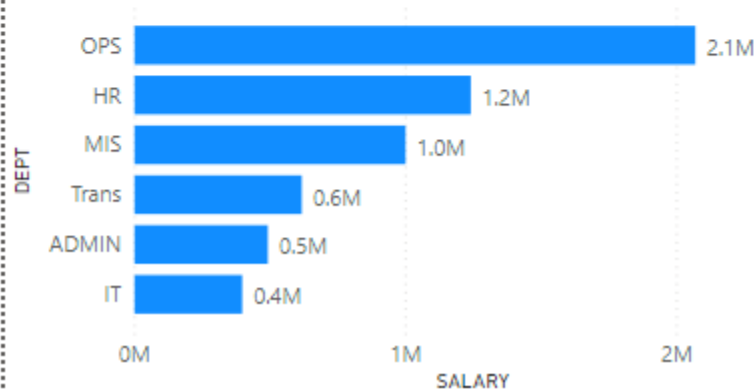
COST ANALYSIS

Training COST by DEPT and Company_Sponsored

Company_Sponsor... ● No ● Yes

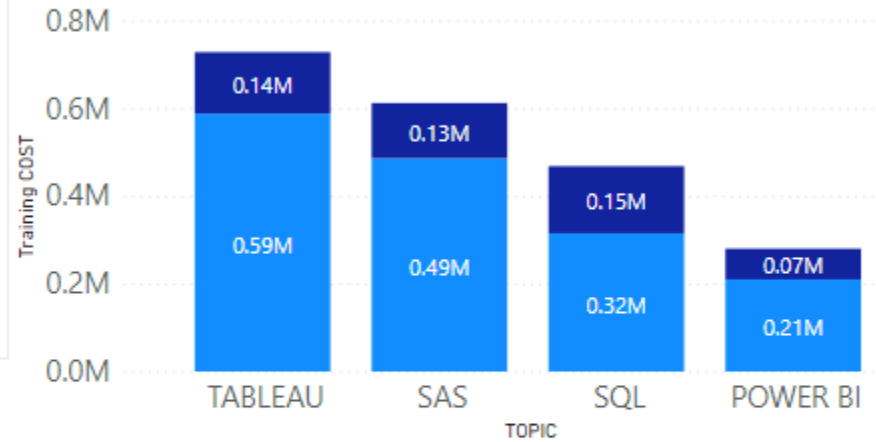


SALARY by DEPT



Training COST and Count of EID by TOPIC and Company_Sponsored

Company_Spon... ● No ● Yes



| TOPIC | Count of EID | CERTIFICATION COST | Training COST |
|----------|--------------|--------------------|---------------|
| POWER BI | 56 | 5000 | 280000 |
| SQL | 52 | 9000 | 468000 |
| TABLEAU | 52 | 14000 | 728000 |
| SAS | 34 | 18000 | 612000 |

