TASK 2

DATA ANALYST INTERNSHIP

EMPLOYEE DATA

Prepared By
Ankush Verma

Data Questions

2

Q1. Can you create a pivot table to summarize the total number of employees in each department?

After creating a Pivot Table, the results reveal the department-wise distribution of the 3000 employees in our dataset. Notably, the Production department boasts the highest employee count, with approximately 2020 individuals, while the Executive Office exhibits the lowest count, with 24 employees. This distribution aligns with expectations and provides valuable insights into the workforce composition across different departments.

Department	Count of Employee
Admin Offices	80
Executive Office	24
IT/IS	430
Production	2020
Sales	331
Software Engineering	115
Grand Total	**3000**

Data Questions

3

Q2. Apply conditional formatting to highlight employees with a "Performance Score" below 3 in red.

The conditional formatting has been successfully applied to the dataset to highlight employees whose performance is below 3.

Following the application of conditional formatting, we discovered that 781 employees fall under the category of underperformance.

A	B	C	D	Z
Employee ID	FirstName	LastName	StartDate	Current Employee Rating
3427	Uriah	Bridges	20-Sep-19	4
3428	Paula	Small	11-Feb-23	3
3429	Edward	Buck	10-Dec-18	4
3430	Michael	Riordan	21-Jun-21	2
3431	Jasmine	Onque	29-Jun-19	3
3432	Maruk	Fraval	17-Jan-20	3
3433	Latia	Costa	6-Apr-22	4
3434	Sharlene	Terry	6-Nov-20	2
3435	Jac	McKinzie	18-Aug-18	3
3436	Joseph	Martins	21-Jan-22	5
3437	Myriam	Givens	4-Aug-23	5
3438	Dheepa	Nguyen	10-Aug-18	3
3439	Bartholemew	Khemmich	25-May-22	3
3440	Xana	Potts	5-Dec-19	3
3441	Prater	Jeremy	28-Apr-19	4
3442	Kaylah	Moon	9-Jul-19	2
3443	Kristen	Tate	5-Apr-21	3
3444	Bobby	Rodgers	28-Nov-21	3
3445	Reid	Park	16-Jan-21	4
3446	Hector	Dalton	24-Aug-21	2
3447	Mariela	Schultz	26-May-20	3

Data Questions

4

Q3. Calculate the average "Satisfaction Score" for male and female employees separately using a pivot table.

First, I utilized the lookup function to retrieve satisfaction scores from the Employee_Engagement dataset and integrate them into the Employee Data. Following the successful acquisition of values for various employees, I employed a Pivot table to display the average satisfaction scores for both males and females.

As observed, the average satisfaction score for males is 3.01, while for females, it is 3.04 out of 5.

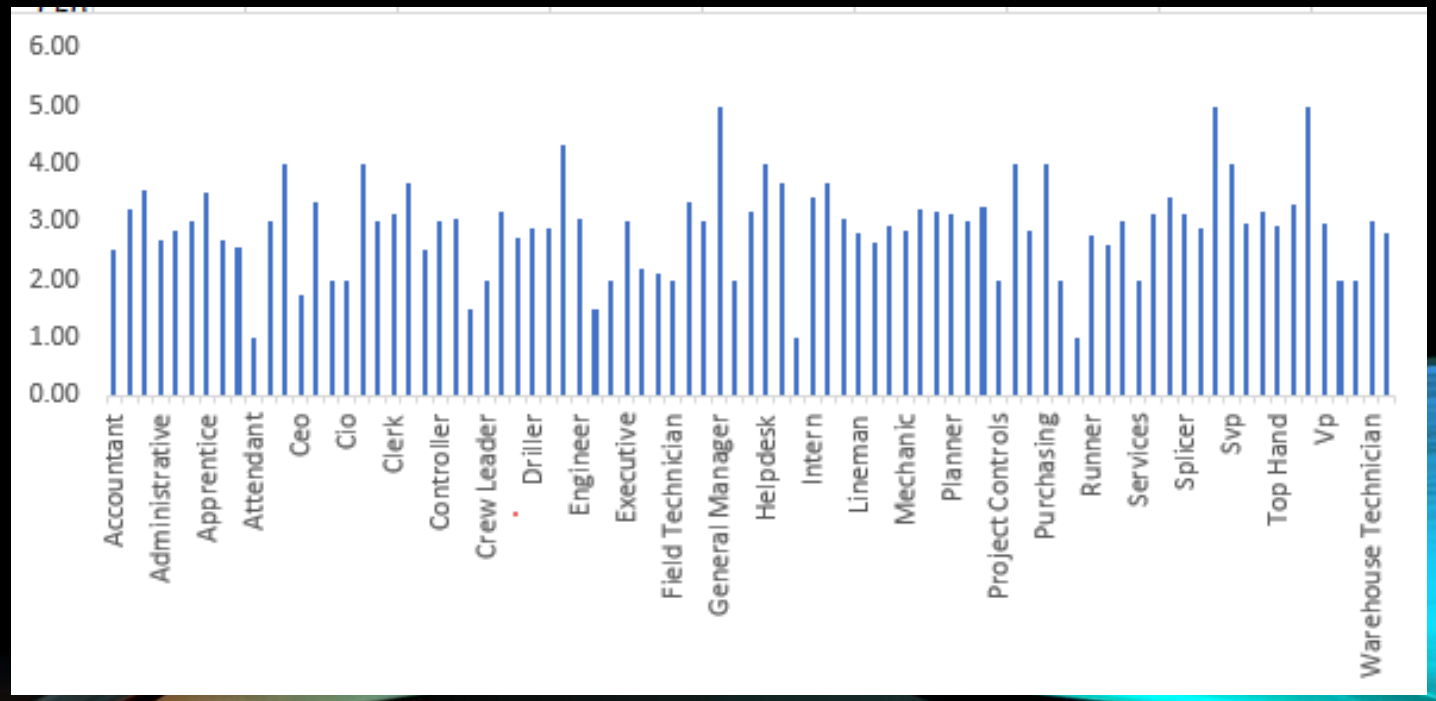
Gender	Average of Satisfaction Score
Female	3.01
Male	3.04
Grand Total	3.02

Data Questions

5

Q4. Create a chart to visualize the distribution of "Work-Life Balance Score" for different job functions.

Here is the snapshot of the column chart that shows the average work-life balance of all employees based on their different job functions.



Data Questions

6

Q5. Filter the data to display only terminated employees and find out the most common "Termination Type."

Termination	Count of TerminationType
Terminated for Cause	**66**
Involuntary	21
Resignation	22
Retirement	10
Voluntary	13
Voluntarily Terminated	**321**
Involuntary	86
Resignation	74
Retirement	76
Voluntary	85
Grand Total	**387**

Data Questions

7

Q6. Calculate the average "Engagement Score" for each department using a pivot table.

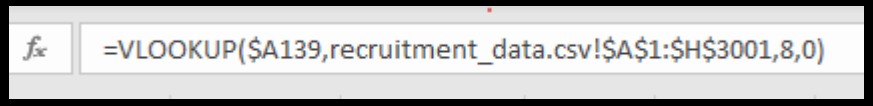
We have successfully created a pivot table that displays the average Engagement Score department-wise for all employees in the employee dataset.

Here, we observe that the Executive Office department has the highest engagement score, which is 3.38 out of 5, while the Production department has the lowest engagement score at 2.91 out of 5.

Department	Average of Engagement Score
Admin Offices	2.93
Executive Office	3.38
IT/IS	3.03
Production	2.91
Sales	2.99
Software Engineering	2.97
Grand Total	**2.94**

Data Questions

Q7. Use VLOOKUP to find the supervisor's email address for a specific employee.



	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1	Sta	JobFunctionDescriptio	GenderCo	LocationCoc	RaceDe	MaritalDe	Performance Sco	Curren	Satis	Wor	Engag	Supervisor Email	
139	MA	Lineman	Female	2122	White	Divorced	Fully Meets	1	4	3	4	rodriguezronald@example.net	
143	MA	Foreman	Male	2478	Asian	Divorced	Fully Meets	2	4	1	2	fharper@example.org	
145	MA	Engineer	Female	2445	Asian	Married	Fully Meets	1	5	3	2	lindasmith@example.net	
146	MA	Engineer	Female	1886	White	Widowed	Fully Meets	4	1	5	4	theresaandrade@example.net	
150	MA	Groundman	Male	2445	Hispanic	Single	Exceeds	5	1	5	5	brooke46@example.org	
173	MA	Supervisor	Female	72750	Other	Married	Needs Improvemen	1	5	4	2	darrellwilliams@example.com	
177	MA	Engineer	Female	30766	Hispanic	Married	Fully Meets	5	3	3	3	matthew78@example.net	
178	MA	Technician	Female	58348	White	Married	Fully Meets	2	1	3	5	qwaller@example.net	
181	MA	Engineer	Male	95618	Other	Divorced	Fully Meets	3	5	1	3	timothynunez@example.net	
200	MA	Engineer	Female	54392	Asian	Single	Fully Meets	5	2	4	5	reyesguy@example.org	
202	MA	Supervisor	Female	52646	Black	Married	Fully Meets	3	1	4	1	alexislong@example.com	
204	MA	Laborer	Male	70056	Black	Divorced	Fully Meets	5	4	4	2	amandasimpson@example.com	
222	MA	Engineer	Female	20847	Black	Single	Fully Meets	3	4	4	2	beardjennifer@example.net	
225	MA	Laborer	Female	72818	White	Divorced	Fully Meets	3	2	4	3	ronald14@example.org	
227	MA	Supervisor	Male	96845	Other	Married	Fully Meets	3	5	2	1	xchen@example.com	
228	MA	Foreman	Male	67277	Asian	Widowed	Exceeds	3	4	3	4	alexanderbryan@example.com	
244	MA	Clerk	Male	81928	Black	Widowed	Fully Meets	3	2	4	4	jerome87@example.com	
246	MA	Splicer	Female	45376	Asian	Divorced	Needs Improvemen	3	4	5	3	carmstrong@example.net	
248	MA	Manager	Female	47767	Other	Widowed	Fully Meets	3	3	1	3	mccoydonna@example.com	
249	MA	Engineer	Female	39141	Other	Single	Fully Meets	3	4	3	4	meganfoster@example.com	

Data Questions

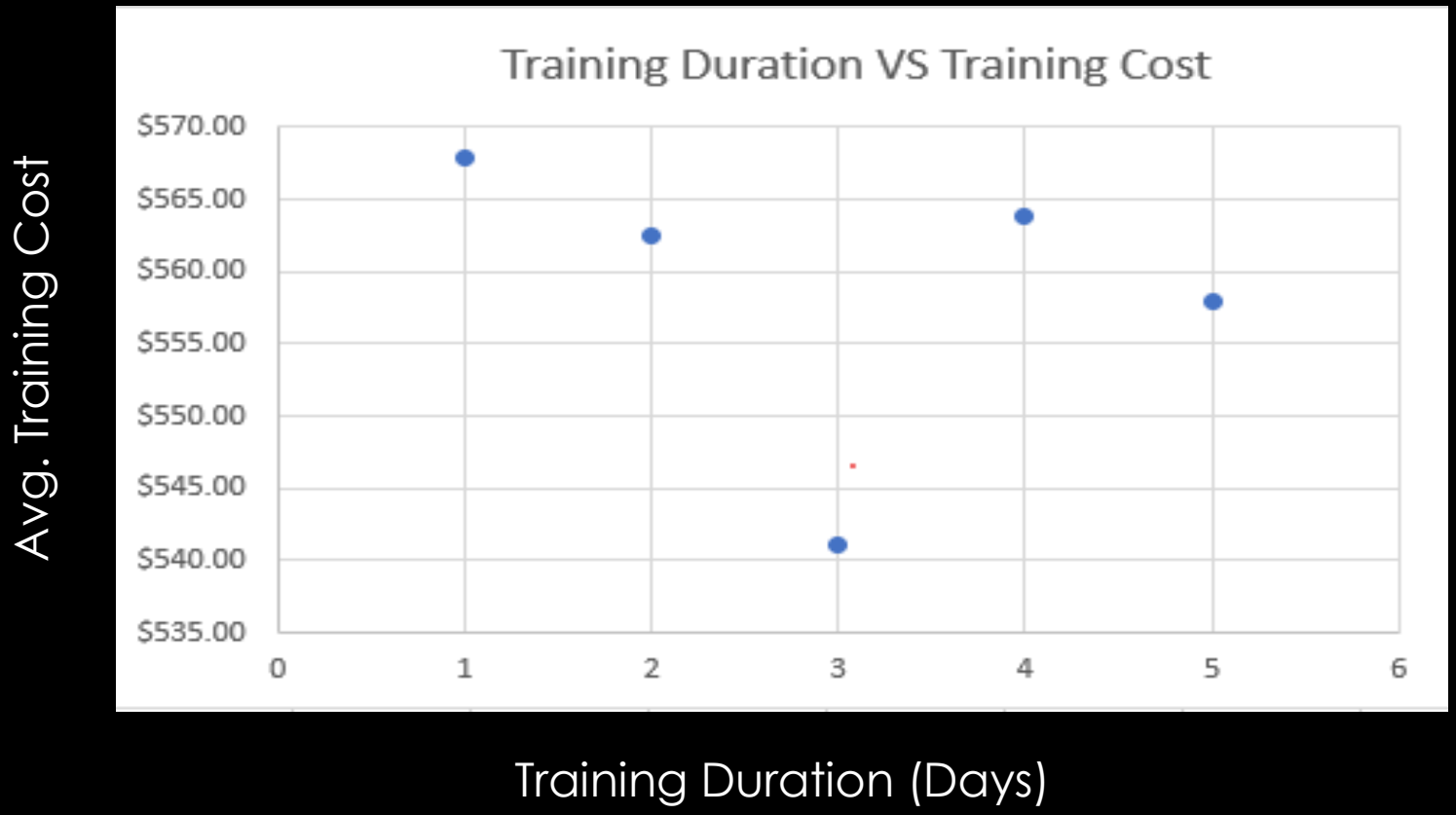
Q8. Can you identify the department with the highest average "Employee Rating?"

After creating a pivot table and sorting avg. employee rating, we have found that the Department of Admin Office has the highest Avg. Employee Rating which is 3.03 out of 5.

Department	Avg. of Employee Rating
Admin Offices	3.03
Production	2.98
IT/IS	2.97
Sales	2.91
Software Engineering	2.90
Executive Office	2.79
Grand Total	**2.97**

Data Questions

Q9. Create a scatter plot to explore the relationship between "Training Duration (Days)" and "Training Cost."



Data Questions

Q10. Build a pivot table that shows the count of employees by "RaceDesc" and "GenderCode."

RaceDesc	Count of Employee ID
Asian	**629**
Female	346
Male	283
Black	**618**
Female	346
Male	272
Hispanic	**572**
Female	325
Male	247
Other	**582**
Female	318
Male	264
White	**599**
Female	347
Male	252
Grand Total	**3000**

Data Questions

Q11. Use INDEX and MATCH functions to find the "Training Program Name" for an employee with a specific ID.

```
=INDEX([training_and_development_data.csv]training_and_development_data!$A$1:$I$3001,MATCH($A2,[training_and_development_data.csv]training_and_development_data!A$1:A$3001,0),MATCH($AD$1,[training_and_development_data.csv]training_and_development_data!A$1:I$1,0))
```

A	B	C	AD
Employee ID	**FirstName**	**LastName**	**Training Program Name**
3427	Uriah	Bridges	Leadership Development
3428	Paula	Small	Customer Service
3429	Edward	Buck	Leadership Development
3430	Michael	Riordan	Project Management
3431	Jasmine	Onque	Technical Skills
3432	Maruk	Fraval	Project Management
3433	Latia	Costa	Customer Service
3434	Sharlene	Terry	Leadership Development
3435	Jac	McKinzie	Customer Service
3436	Joseph	Martins	Leadership Development
3437	Myriam	Givens	Technical Skills
3438	Dheepa	Nguyen	Technical Skills
3439	Bartholemew	Khemmich	Technical Skills
3440	Xana	Potts	Communication Skills
3441	Prater	Jeremy	Project Management
3442	Kaylah	Moon	Project Management
3443	Kristen	Tate	Technical Skills
3444	Bobby	Rodgers	Technical Skills
3445	Reid	Park	Project Management
3446	Hector	Dalton	Technical Skills

Data Questions

Q12. Create a multi-level pivot table to analyze the "Performance Score" by "BusinessUnit" and "JobFunctionDescription."

We have successfully created a multilevel pivot table showing the "performance score" as a count for different "Business Unit" and "JobFunctionDescription."

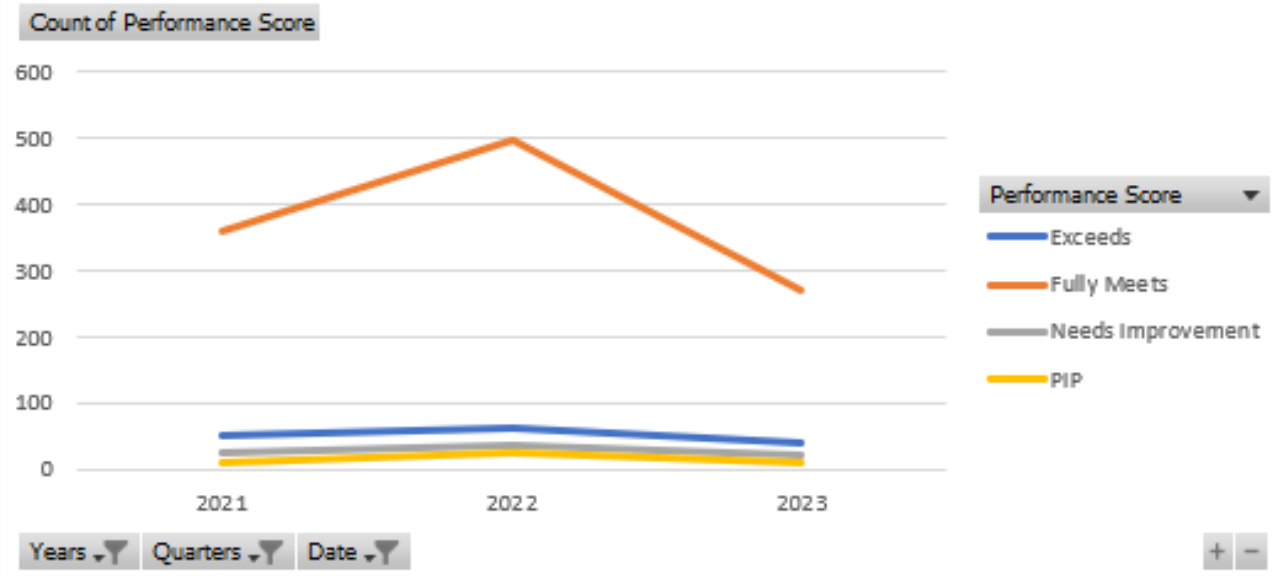
If we expand the rows, i.e., "Business Unit," by clicking on the "+" sign, we can see the count of "performance scores" for specific "JobFunctionDescription."

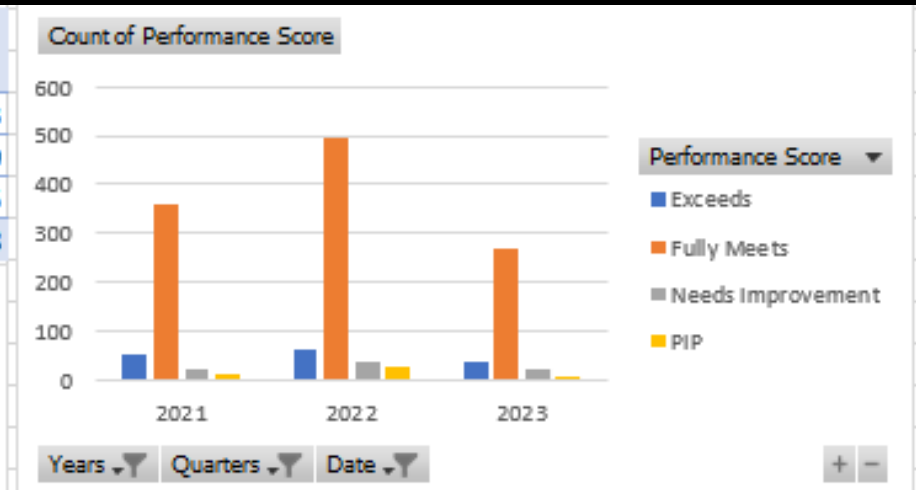
Business Unit	Count of Performance Score
⊞ BPC	303
⊞ CCDR	300
⊞ EW	302
⊞ MSC	296
⊞ NEL	304
⊞ PL	301
⊞ PYZ	299
⊞ SVG	304
⊞ TNS	297
⊞ WBL	294
Grand Total	3000

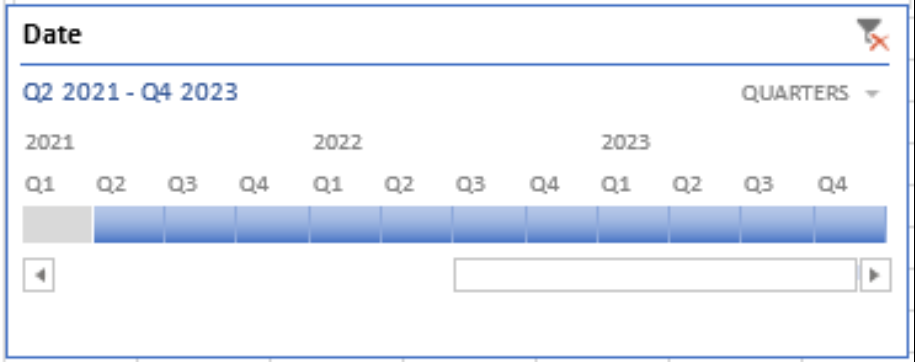
Data Questions

Q13. Design a dynamic chart that allows users to select and visualize the performance of any employee over time.

Count of Performance Score	Column Labels				
DATE	Exceeds	Fully Meets	Needs Improvement	PIP	Grand Total
2021	50	359	24	10	443
2022	62	497	36	25	620
2023	38	269	20	8	335
Grand Total	150	1125	80	43	1398



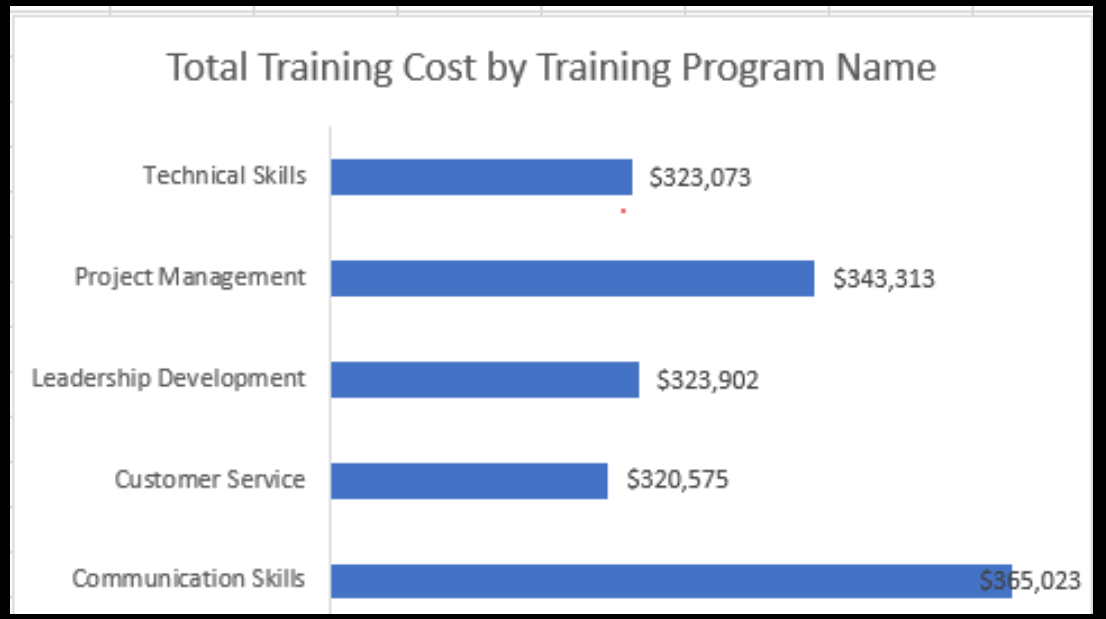




Data Questions

Q14. Calculate the total training cost for each "Training Program Name" and display it in a bar chart.

Training Program	Sum of Training Cost
Communication Skills	$365,023
Customer Service	$320,575
Leadership Development	$323,902
Project Management	$343,313
Technical Skills	$323,073
Grand Total	**$1,675,886**



Data Questions

Q15. Apply advanced conditional formatting to highlight the top 10% and bottom 10% of employees based on "Current Employee Rating."

JobFunctionDescription	GenderCode	LocationCode	RaceDesc	MaritalDesc	Performance Score	Current Employee Rating	Sati
Accounting	Female	34904	White	Widowed	Fully Meets	4	
Labor	Male	6593	Hispanic	Widowed	Fully Meets	3	
Assistant	Male	2330	Hispanic	Widowed	Fully Meets	4	
Clerk	Male	58782	Other	Single	Fully Meets	2	
Laborer	Female	33174	Other	Married	Fully Meets	3	
Driver	Male	6050	Black	Married	Fully Meets	3	
Technician	Female	90007	Hispanic	Divorced	Exceeds	4	
Engineer	Female	97756	White	Divorced	Fully Meets	2	
Executive Assistant	Male	78789	Black	Widowed	Exceeds	3	
Engineer	Male	78207	Asian	Widowed	Fully Meets	5	
Technician	Female	46204	Other	Single	Fully Meets	5	
Technician	Female	30428	Asian	Married	Fully Meets	3	
Splicer	Male	80820	Other	Single	Fully Meets	3	
Controller	Female	40220	White	Divorced	Fully Meets	3	
Lineman	Male	89139	Asian	Widowed	Exceeds	4	
Laborer	Male	2810	Black	Single	Exceeds	2	
Coordinator	Male	2621	Asian	Widowed	Fully Meets	3	
Director	Male	44553	Other	Widowed	Fully Meets	3	
Supervisor	Female	5360	Other	Married	Exceeds	4	
Driller	Female	16325	White	Divorced	Exceeds	2	

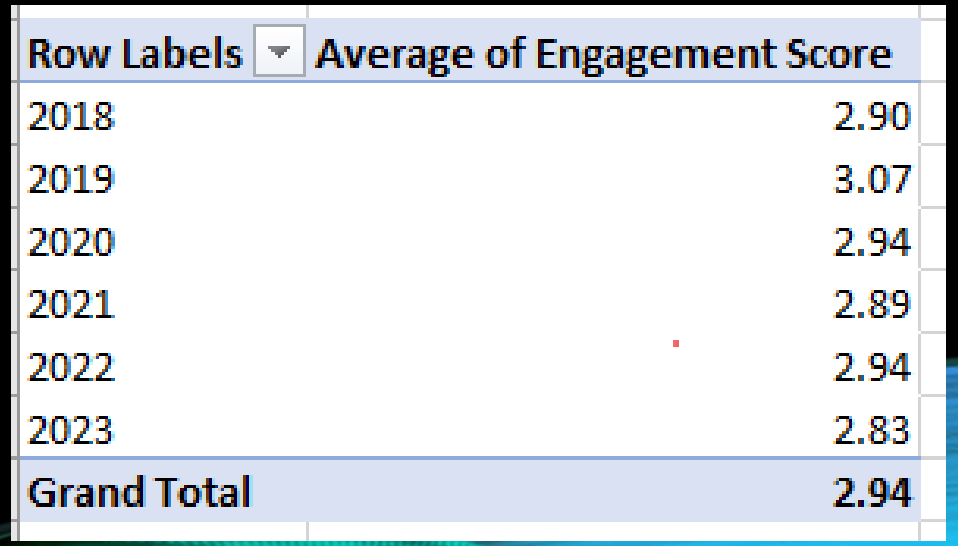
Data Questions

Q16. Use a calculated field in a pivot table to determine the average "Engagement Score" per year.

While Calculated Fields in PivotTables can utilize mathematical operations and some functions, they have a limitation when it comes to directly applying aggregate functions like Average or Sum within their formulas.

However, you can effectively achieve the same result by:

- Adding the field you want to average (e.g., Engagement Score) to the Values area of the PivotTable.
- Right-clicking on the field and selecting "Value Field Settings."
- Under "Summarize Values By," choosing "Average."
- This will display the average of the field's values within the
- PivotTable, even though it's not directly calculated within a Calculated Field.



Row Labels	Average of Engagement Score
2018	2.90
2019	3.07
2020	2.94
2021	2.89
2022	2.94
2023	2.83
Grand Total	2.94

Data Questions

Q17. Can you build a macro that automates the process of updating and refreshing all pivot tables in the workbook?

- To insert VBA Macro, we first need to Enable the Developer tab:
1) Go to File > Options > Customize Ribbon.
2) Under "Main Tabs," check the box for "Developer."

- Open the Visual Basic Editor:
1) Click on the Developer tab and then click "Visual Basic."

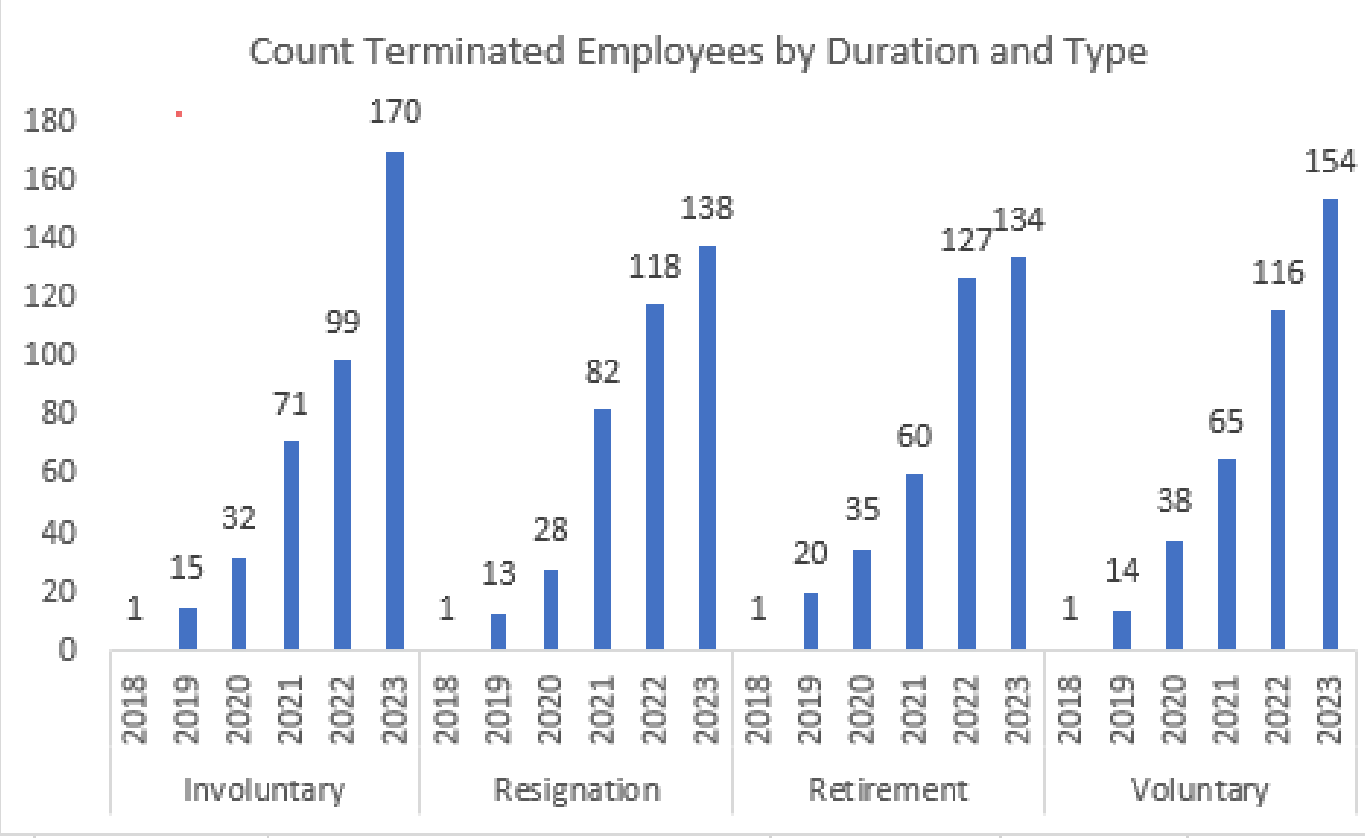
- Insert the code:
1) In the VBA editor, double-click on the workbook name in the Project Explorer.
2) Paste the provided code into the code window.

- Run the macro:
1) Close the VBA editor.
2) Go back to your Excel workbook.
3) Press Alt + F8 to open the Macros dialog box.
4) Select "RefreshAllPivotTables" and click "Run."
This will automatically refresh all pivot tables in your workbook.

```
Sub RefreshAllPivotTables()

    Dim pt As PivotTable

    For Each pt In ActiveWorkbook.PivotTables
        pt.RefreshTable
    Next pt

End Sub
```

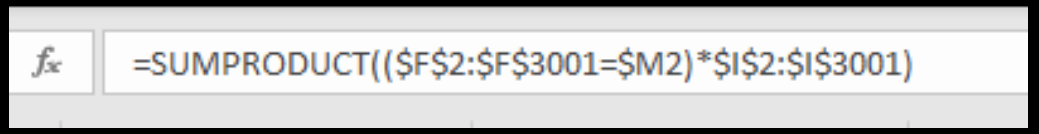
Data Questions

Q18. Create a histogram to understand the distribution of "ExitDate" for terminated employees.



Data Questions

Q19. Utilize the SUMPRODUCT function to calculate the total training cost for employees in a specific location.



M	N
Location	**Training Cost**
Port Greg	510.83
Brandonview	1108.45
Port Briannahaven	777.06
Knightborough	824.3
Bruceshire	145.99
Erinfort	838.07
New Christopher	667.32
Lowemouth	758.18

Data Questions

Q20. Develop a dashboard that provides an overview of key HR metrics, including headcount, performance, and training costs, using charts and pivot tables.