5/4/2024

Employee Data Analysis

Internship Project I

MOGHEES HASHMI
DATA ANALYST INTERNAT PSYLIQ

ABSTRACT

The Employee Data Analysis project undertaken during the Data Analyst Internship at **Psyliq** involved leveraging tools such as Power BI, Excel, and PowerPoint to analyse and derive insights from employee data. The project aimed to provide actionable insights to enhance **organizational performance**, **employee satisfaction**, and **decision-making processes**.

Using **Power BI**, the project focused on creating interactive dashboards and visualizations to explore various aspects of employee data, including demographics, performance metrics, training history, and turnover rates. **Excel** was utilized for data preprocessing, cleaning, and **advanced analytics**, enabling deep dives into complex datasets and **statistical analyses**.

The project encompassed multiple phases, including data **collection**, **cleaning**, **analysis**, **visualization**, and **presentation.**

By Moghees Hashmi

Date: 04-05-2024

Project GitHub Link:
https://github.com/hashmi7238/Employee_Data_Analysis_Intern_Project

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

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

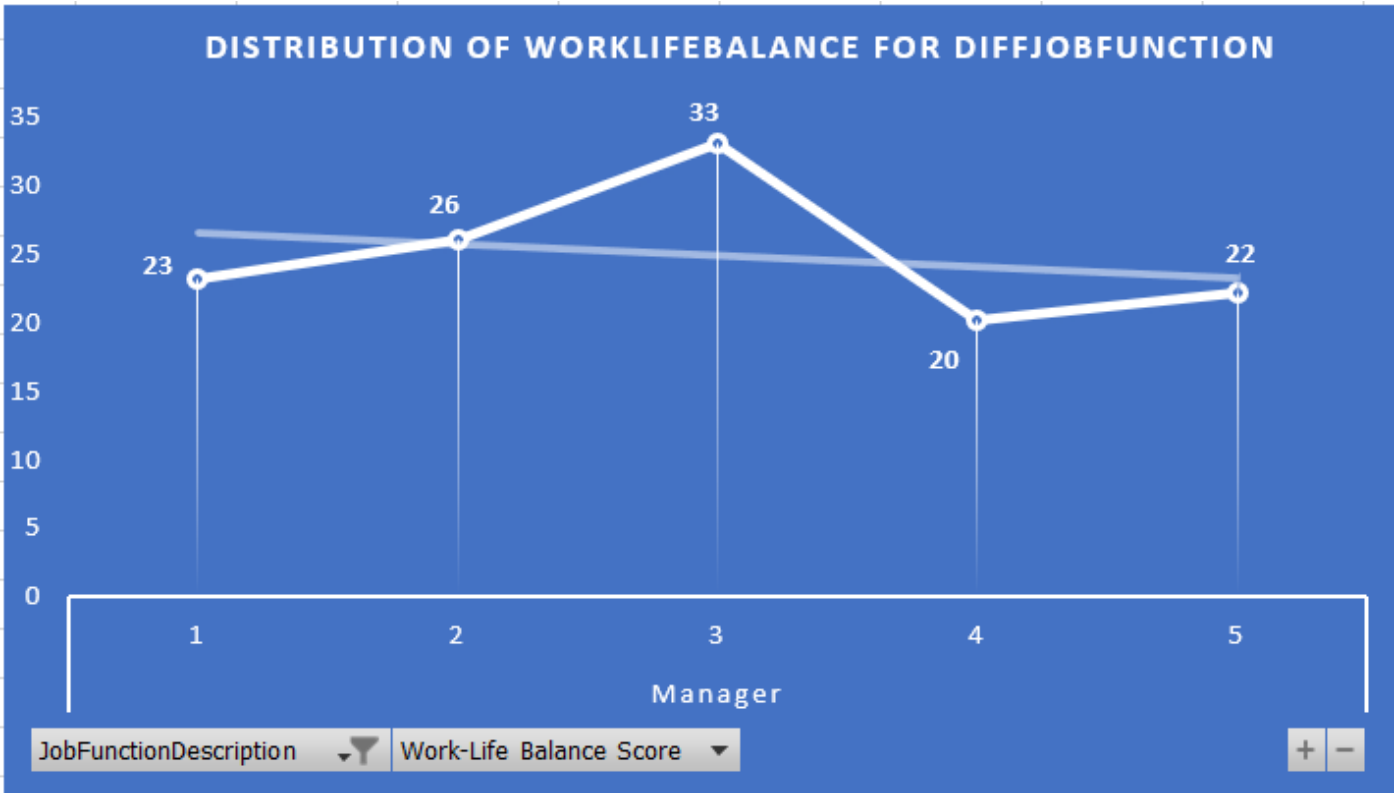
2. Apply conditional formatting to highlight employees with a "Performance Score" below 3 in red.

Employee ID	FirstName	LastName	StartDate	ExitDate	Title	Supervisor
3451	Carlee	French	18-Feb-21	11-Nov-22	Area Sales Manager	Michael Bradshaw
3452	Jaydon	Blackburn	08-Nov-22		Area Sales Manager	Debbie Crosby
3453	Bridger	Carter	13-Oct-22		Area Sales Manager	Elizabeth Taylor
3454	Leon	Beard	11-Sep-22		Area Sales Manager	Brian Dunlap
3455	Charity	Miranda	29-Jun-21	05-Jul-22	Area Sales Manager	Daniel Rodriguez
3456	Axel	Howe	06-Mar-23		Area Sales Manager	Martin Hammond
3457	Milton	Wall	25-Sep-20		Area Sales Manager	Sara Bell
3458	Cory	Robinson	28-Apr-22	24-May-23	Area Sales Manager	Elizabeth Anderson
3459	Saniya	Yu	18-Apr-21	21-Jun-22	Area Sales Manager	Erin Bailey
3460	Alisa	James	19-Feb-20		Area Sales Manager	Dennis Henderson
3461	Lincoln	Compton	18-Jul-19	01-Oct-21	Area Sales Manager	Tammy Conner
3462	Aliana	Nolan	13-Sep-18	17-Sep-21	Area Sales Manager	Brianna Mathews
3463	Kayden	Dodson	30-Apr-20	18-Jul-21	Area Sales Manager	Jessica Peters
3464	James	Duke	03-Dec-18	09-Sep-22	Area Sales Manager	Mark Harris
3465	Willow	Stuart	08-Nov-21		Area Sales Manager	Elizabeth Tapia
3466	Clayton	Walker	13-Apr-22	10-Apr-23	Area Sales Manager	Jon Holden
3467	Celia	Curtis	06-May-20		Area Sales Manager	Michael Odonnell
3468	Valentin	Reilly	09-Sep-19		Area Sales Manager	Andres Parker
3469	Ryland	Shepherd	29-Jul-20	05-Jan-23	Area Sales Manager	James Berry
3470	Esteban	Gilbert	14-Nov-18	28-Oct-19	Area Sales Manager	Gail Ali

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

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

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



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

>>> Most termination type is "**Involuntary**" with 107 terminations.

Termination Type	Count of TerminationType
Involuntary	107
Resignation	96
Retirement	86
Voluntary	98
Grand Total	**387**

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

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**

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

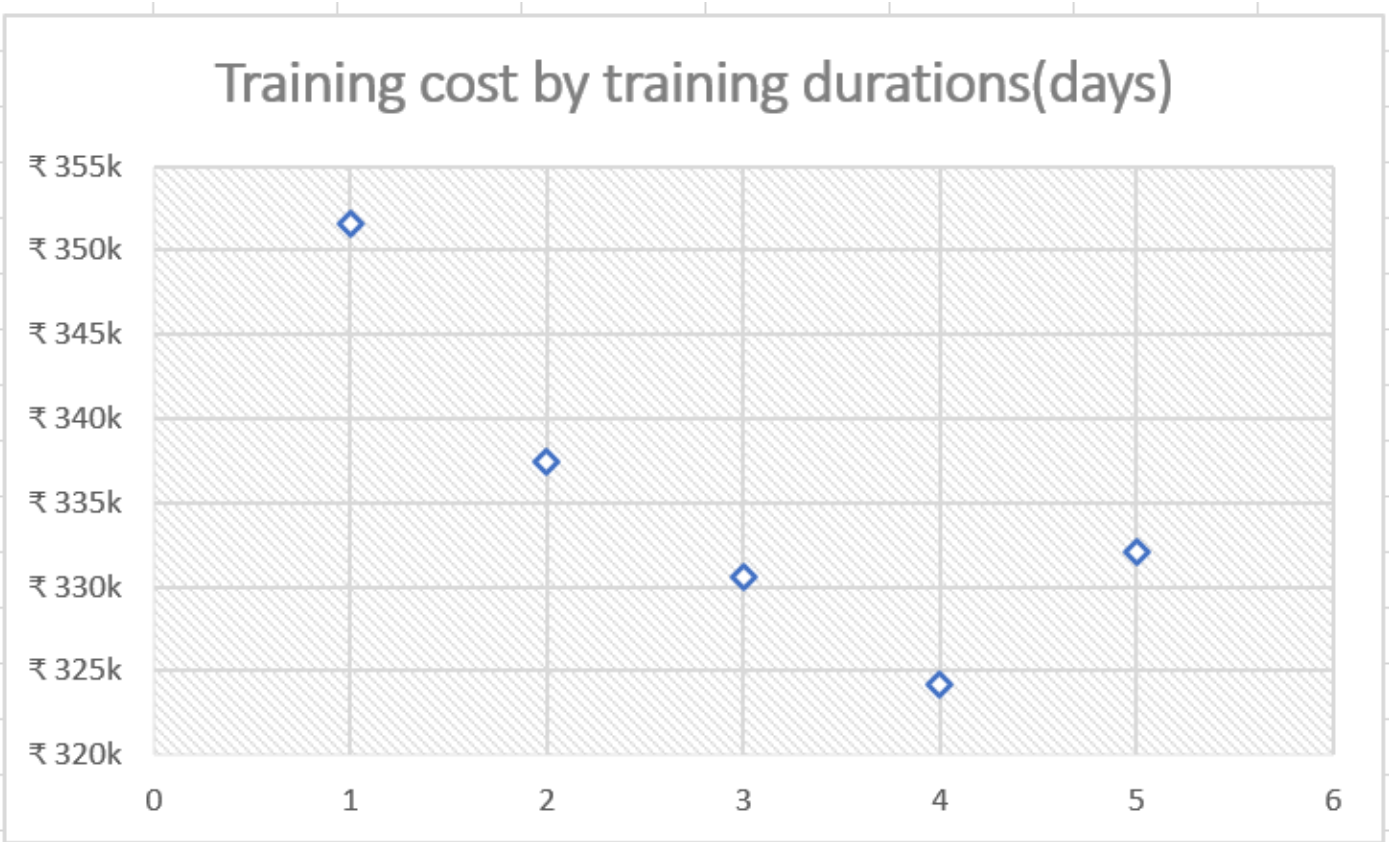
Employee ID	Email
3430	=VLOOKUP(AB2,$A$1:$H$3001,8,FALSE)

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

>>> **Admin Offices** has average employee rating.

Department	Average of Current Employee Rating
Admin Offices	3.03
Executive Office	2.79
IT/IS	2.97
Production	2.98
Sales	2.91
Software Engineering	2.90

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



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

Count of Employee	Gender		
Race	Female	Male	Grand Total
Asian	346	283	629
Black	346	272	618
Hispanic	325	247	572
Other	318	264	582
White	347	252	599
Grand Total	1682	1318	3000

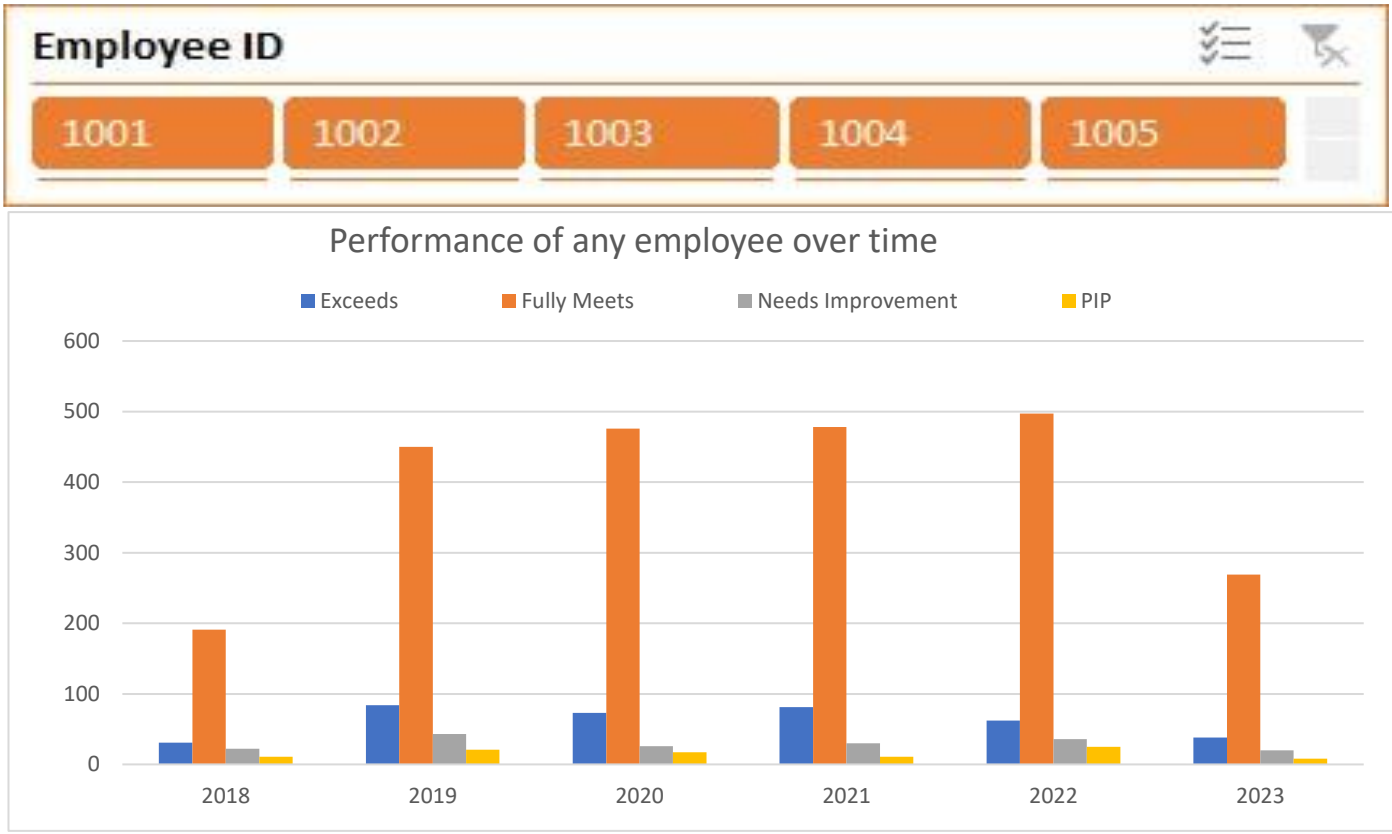
11. Use INDEX and MATCH functions to find the "Training Program Name" for an employee with a specific ID.

Employee ID	Training Program Name
1001	=INDEX(A2:C3001,MATCH(K2,A2:A3001,0),3)

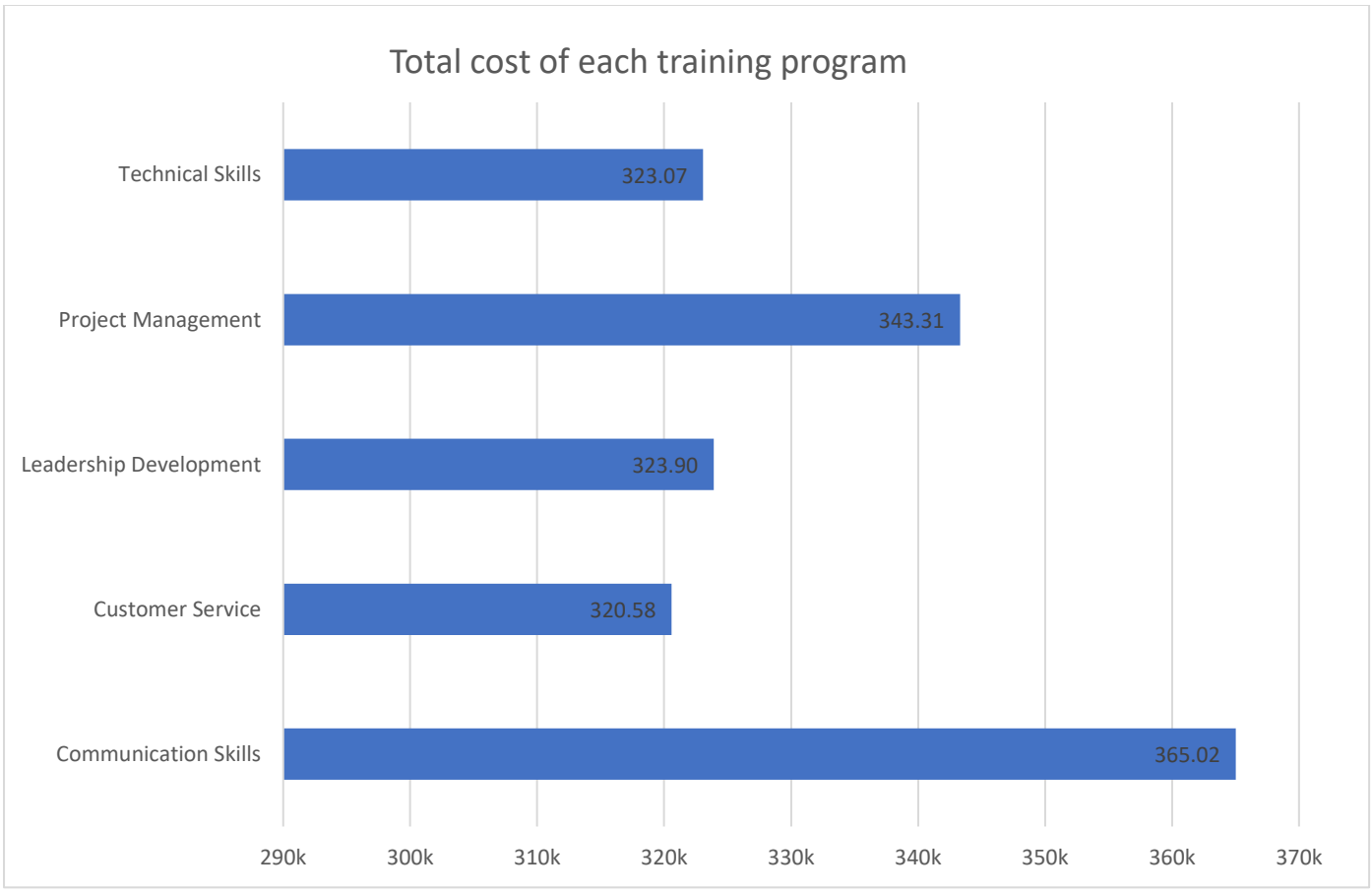
12. Create a multi-level pivot table to analyse the "Performance Score" by "BusinessUnit" and "JobFunctionDescription."

Count of Performance Score	Business Unit										
JobFunctionDescription	BPC	CCDR	EW	MSC	NEL	PL	PYZ	SVG	TNS	WBL	Grand Total
Manager	15	16	9	14	10	11	17	9	14	9	124
Accountant										2	2
Accounting		1		1	1		2	1	3		9
Administration	3	1	1	6	1	1	2	3	2	3	23
Administrative	7	4	7	10	5	4	2	5	5	7	56
Administrator	9	3	2	5	2	6	6	7	4	6	50
Analyst						1					1
Apprentice		1	1		1			1	1	1	6
Assistant	1	5	3	3	4	3	4		1		24
Associate	3	1		1	1		1				7
Attendant						1					1
Billing	2		2		1	1	1	1	3	2	13
Business Development									1		1
Ceo	1						1	1	1		4
Cfo				1		2					3
Chief Operating Officer					1						1
Cio	1										1

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



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



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

Employee ID	FirstName	LastName	StartDate	Year	ExitDate	Title	Supervisor
3427	Uriah	Bridges	20-Sep-19	2019		Production Technician I	Peter Oneill
3428	Paula	Small	11-Feb-23	2023		Production Technician I	Renee Mccormick
3429	Edward	Buck	10-Dec-18	2018		Area Sales Manager	Crystal Walker
3430	Michael	Riordan	21-Jun-21	2021		Area Sales Manager	Rebekah Wright
3431	Jasmine	Onque	29-Jun-19	2019		Area Sales Manager	Jason Kim
3432	Maruk	Fraval	17-Jan-20	2020		Area Sales Manager	Sheri Campos
3433	Latia	Costa	06-Apr-22	2022	03-Jul-23	Area Sales Manager	Jacob Braun
3434	Sharlene	Terry	06-Nov-20	2020	29-Jan-23	Area Sales Manager	Tracy Marquez
3435	Jac	McKinzie	18-Aug-18	2018		Area Sales Manager	Sharon Becker
3436	Joseph	Martins	21-Jan-22	2022	29-Jun-23	Area Sales Manager	George Jenkins
3437	Myriam	Givens	04-Aug-23	2023		Area Sales Manager	Troy White
3438	Dheepa	Nguyen	10-Aug-18	2018	04-Nov-19	Area Sales Manager	Brian Miller
3439	Bartholemew	Khemmich	25-May-22	2022	27-Nov-22	Area Sales Manager	Charles Parks
3440	Xana	Potts	05-Dec-19	2019	17-Feb-23	Area Sales Manager	Gregory Walker
3441	Prater	Jeremy	28-Apr-19	2019		Area Sales Manager	Tyler Lewis
3442	Kaylah	Moon	09-Jul-19	2019	16-Jun-22	Area Sales Manager	Ashley Scott
3443	Kristen	Tate	05-Apr-21	2021	12-May-23	Area Sales Manager	Lauren Jones
3444	Bobby	Rodgers	28-Nov-21	2021	04-Feb-22	Area Sales Manager	Matthew Jackson
3445	Reid	Park	16-Jan-21	2021		Area Sales Manager	Michelle Mitchell
3446	Hector	Dalton	24-Aug-21	2021		Area Sales Manager	Sydney French
3447	Mariela	Schultz	26-May-20	2020	18-Jun-23	Area Sales Manager	Michelle Evans MD
3448	Angela	Molina	01-Oct-19	2019	06-Nov-20	Area Sales Manager	Patricia Cook

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

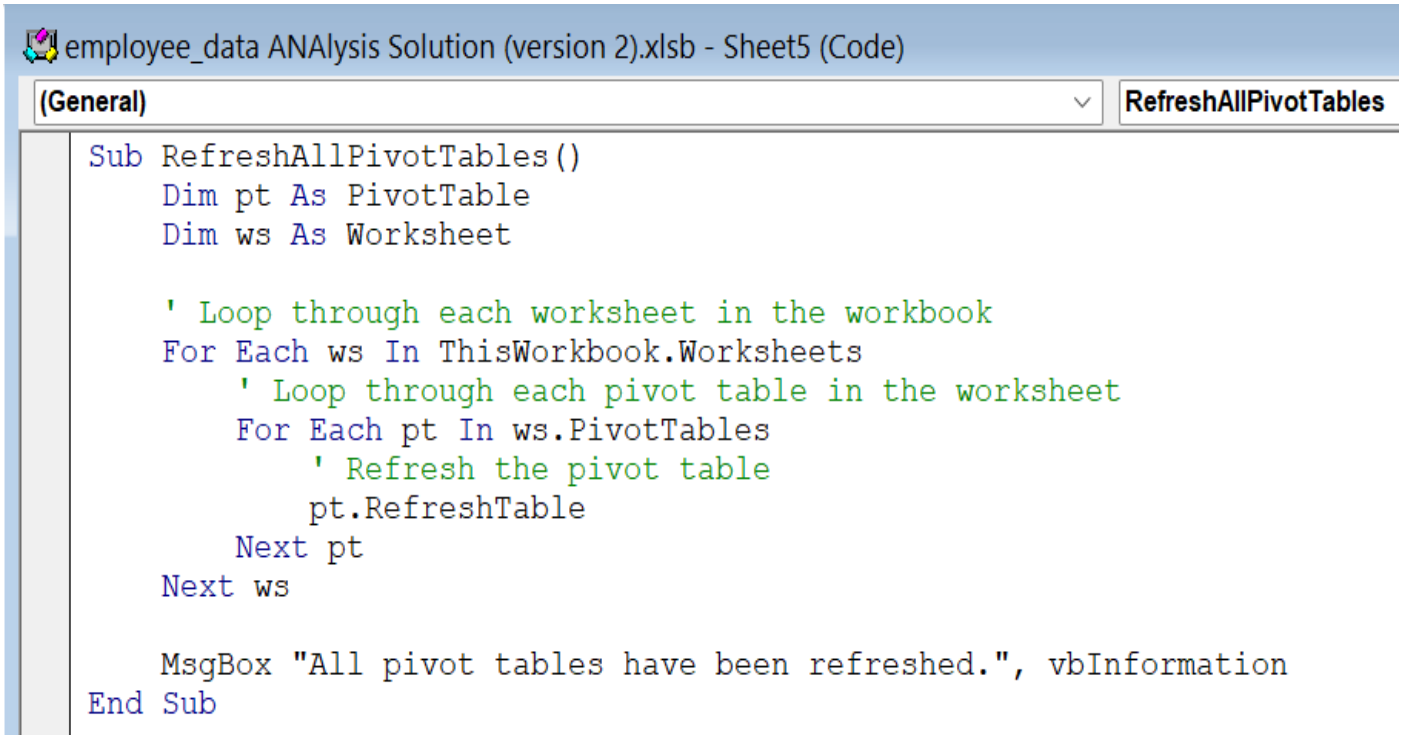
Year	Average of Engagement Score
2018	2.90
2019	3.07
2020	2.94
2021	2.89
2022	2.94
2023	2.83

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

>>Open VBA Editor ALT + F11

>>Write this code in module

>>Close Editor and run micros ALT + F8

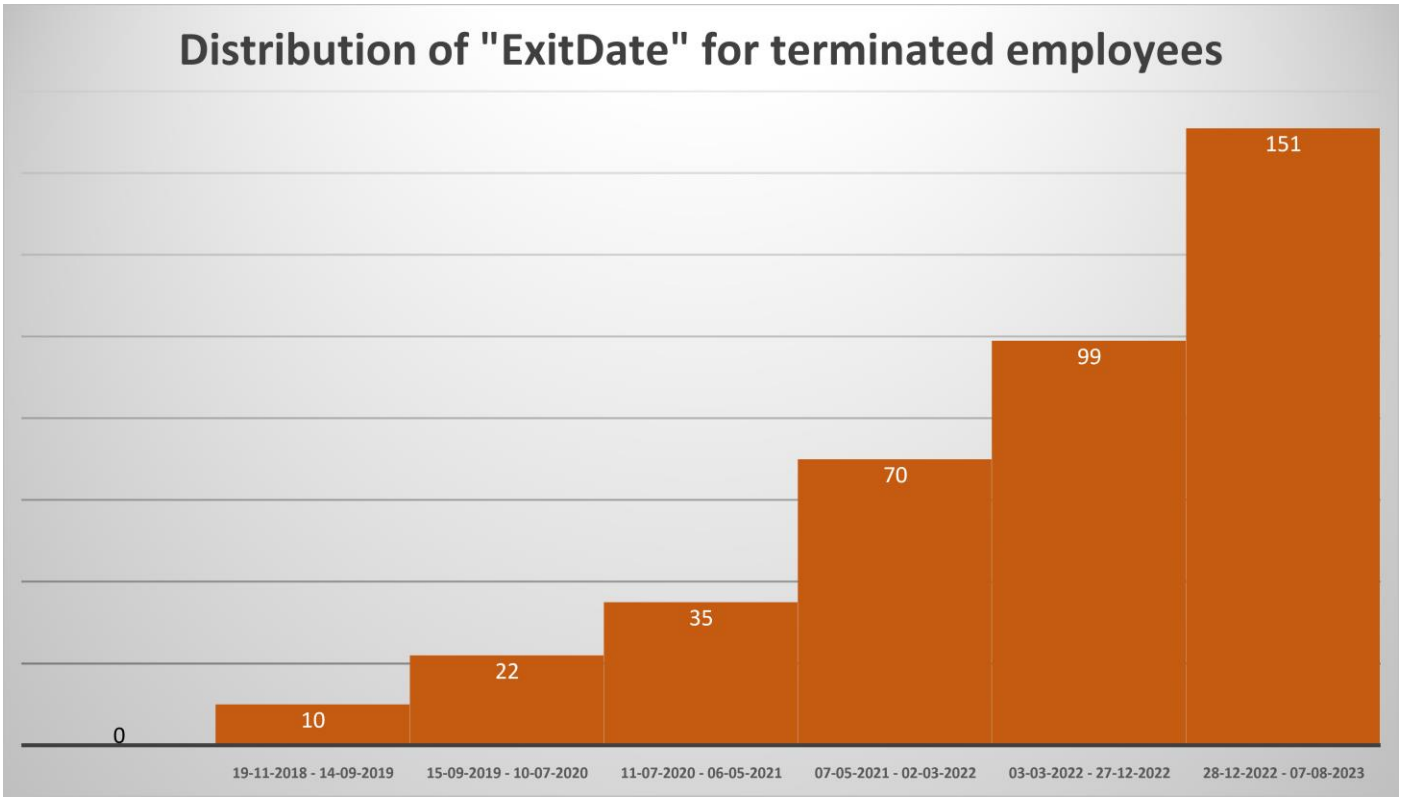



```
Sub RefreshAllPivotTables()
    Dim pt As PivotTable
    Dim ws As Worksheet

    ' Loop through each worksheet in the workbook
    For Each ws In ThisWorkbook.Worksheets
        ' Loop through each pivot table in the worksheet
        For Each pt In ws.PivotTables
            ' Refresh the pivot table
            pt.RefreshTable
        Next pt
    Next ws

    MsgBox "All pivot tables have been refreshed.", vbInformation
End Sub
```

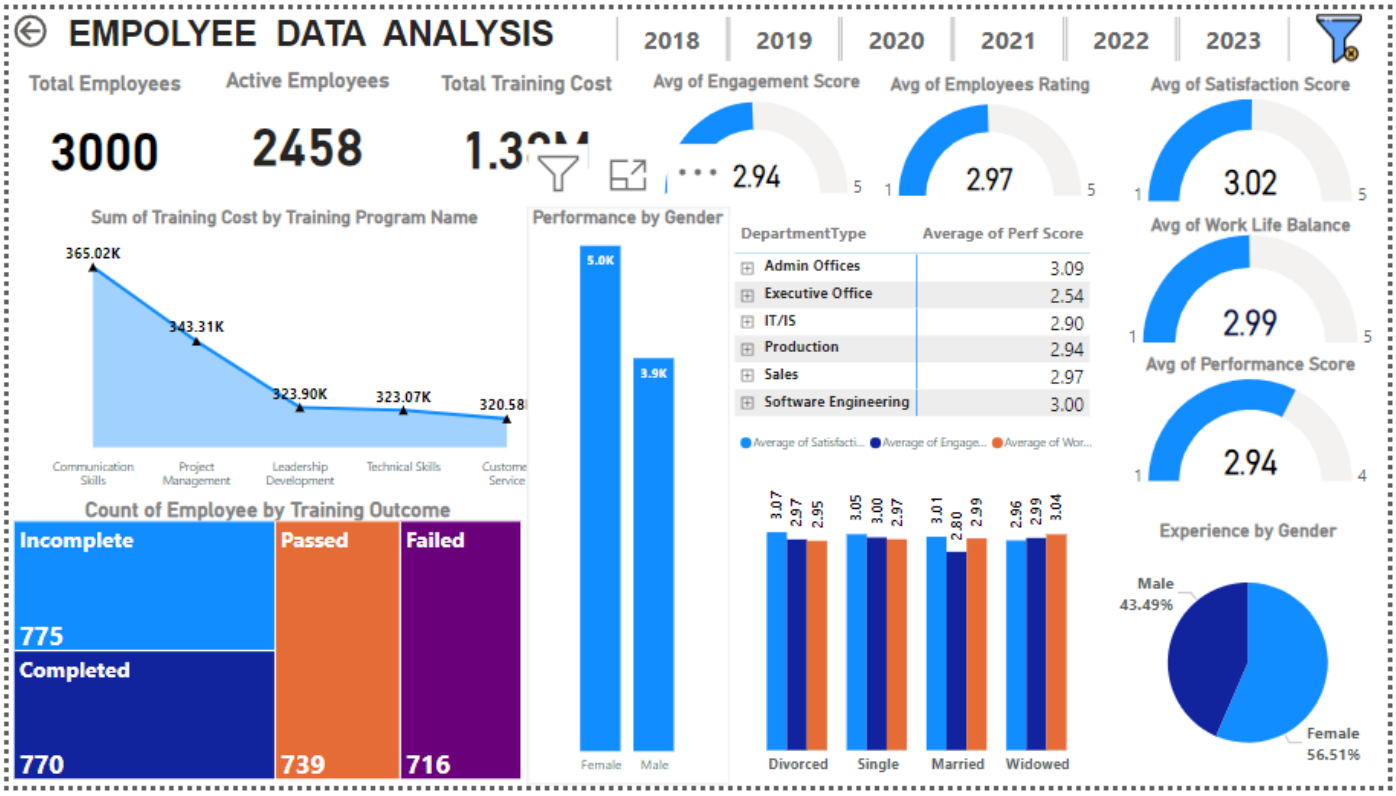
18. Create a histogram to understand the distribution of "ExitDate" for terminated employees.



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

Location	Total Cost
New Ashley	=SUMPRODUCT((F2:F3001=K5)*(I2:I3001))

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



LinkedIn: https://www.linkedin.com/in/moghees-hashmi/

GitHub: https://github.com/hashmi7238

Portfolio: https://hashmi7238.github.io/hashmi-portfolio.github.oi/