DATA ANALYST INTERNSHIP

TASK - 2
EMPLOYEE DATA ANALYSIS

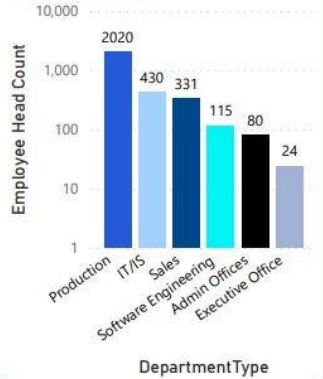
BY
SWARNAV KUMAR
Email id :- swarnav786@gmail.com

EMPLOYEE DATA ANALYSIS

Gender: All

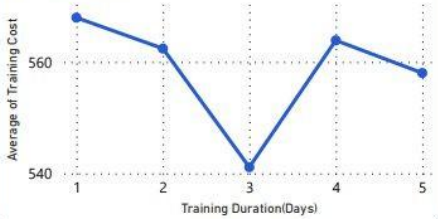
City: All

Employee Head Count by Department Type:



3000	2458	65.08K
Employee Count	Active Employees	Avg. Desired Salary

Average of Training Cost by Training Duration(Days):

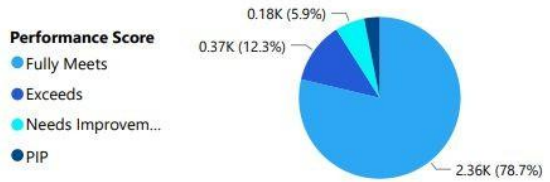


100.04	999.97
Min. Training Cost	Max. Training Cost

State and Employee Classification Type:

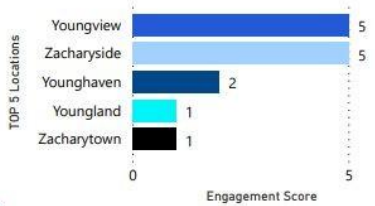
Training Program Name	Sum of Training Cost
Communication Skills	3,65,023.24
Customer Service	3,20,575.04
Leadership Development	3,23,902.03
Project Management	3,43,313.17
Technical Skills	3,23,072.61
Total	**16,75,886.09**

Employee Count by Performance Score:



Total Satisfaction Score by Business Unit:

Engagement Score by TOP 5 Locations:



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

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

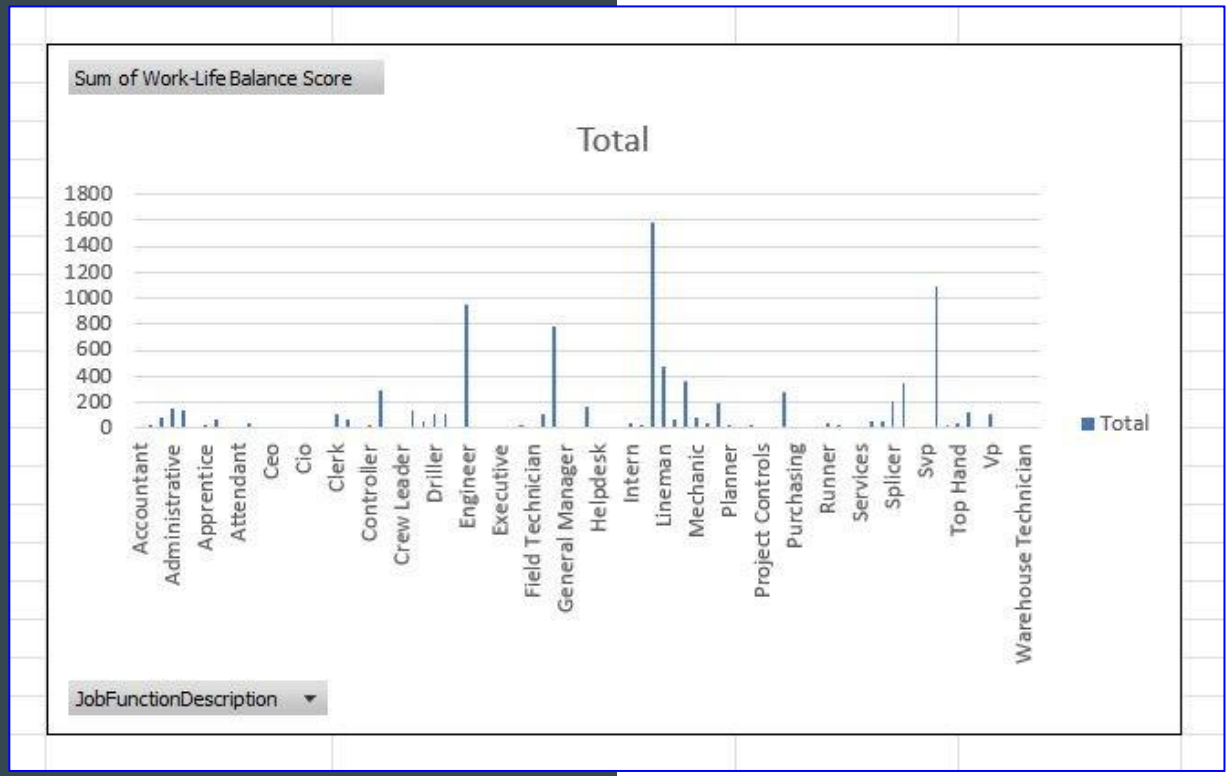
Apply conditional formatting to highlight employees with a "Current employee rating" below 3 in red.

					CER < 3
GenderCode	LocationCode	RaceDesc	MaritalDesc	Performance Score	Current Employee Rating
Female	34904	White	Widowed	Fully Meets	4
Male	6593	Hispanic	Widowed	Fully Meets	3
Male	2330	Hispanic	Widowed	Fully Meets	4
Male	58782	Other	Single	Fully Meets	2
Female	33174	Other	Married	Fully Meets	3
Male	6050	Black	Married	Fully Meets	3
Female	90007	Hispanic	Divorced	Exceeds	4
Female	97756	White	Divorced	Fully Meets	2
Male	78789	Black	Widowed	Exceeds	3
Male	78207	Asian	Widowed	Fully Meets	5
Female	46204	Other	Single	Fully Meets	5
Female	30428	Asian	Married	Fully Meets	3
Male	80820	Other	Single	Fully Meets	3
Female	40220	White	Divorced	Fully Meets	3
Male	89139	Asian	Widowed	Exceeds	4
Male	2810	Black	Single	Exceeds	2
Male	2621	Asian	Widowed	Fully Meets	3
Male	44553	Other	Widowed	Fully Meets	3
Female	5360	Other	Married	Exceeds	4
Female	16325	White	Divorced	Exceeds	2
Female	43481	Asian	Widowed	Fully Meets	3
Male	50705	Asian	Widowed	Fully Meets	3
Male	5168	Black	Single	Fully Meets	5
Male	11765	Other	Single	Exceeds	2
Male	71071	Other	Divorced	Exceeds	3
Male	12122	Black	Divorced	Fully Meets	2
Male	87065	White	Widowed	Fully Meets	4
Male	10415	Hispanic	Married	Needs Improvement	4

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

Row Labels	Average of Satisfaction Score
Female	3.005945303
Male	3.042488619
Grand Total	3.022

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



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

Row Labels	Count of Employee ID
Involuntary	388
Resignation	380
Retirement	377
Unk	1467
Voluntary	388
Grand Total	**3000**

According to this table, "Unk" is the most common Termination type.

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

Row Labels	Average of Engagement Score
Admin Offices	2.925
Executive Office	3.375
IT/IS	3.025581395
Production	2.906435644
Sales	2.990936556
Software Engineering	2.973913043
Grand Total	**2.939666667**

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

Employee ID	Supervisor	Email Id
1001	Angela Carlson	susan.exantus@bilearner.com
1002	Angela Hayes	sandra.martin@bilearner.com
1003	Christina Copeland	keyla.del bosque@bilearner.com
1004	Jennifer Cohen	andrew.szabo@bilearner.com
1005	Mr. Jesus Richards	luke.patronick@bilearner.com
1006	Beth Johnson	colby.andreola@bilearner.com
1007	Raymond Adams	edward.true@bilearner.com
1008	Jessica Rhodes	judith.carabbio@bilearner.com
1009	Steven Tran	adell.saada@bilearner.com
1010	Kristen Collier	kamari.hunter@bilearner.com

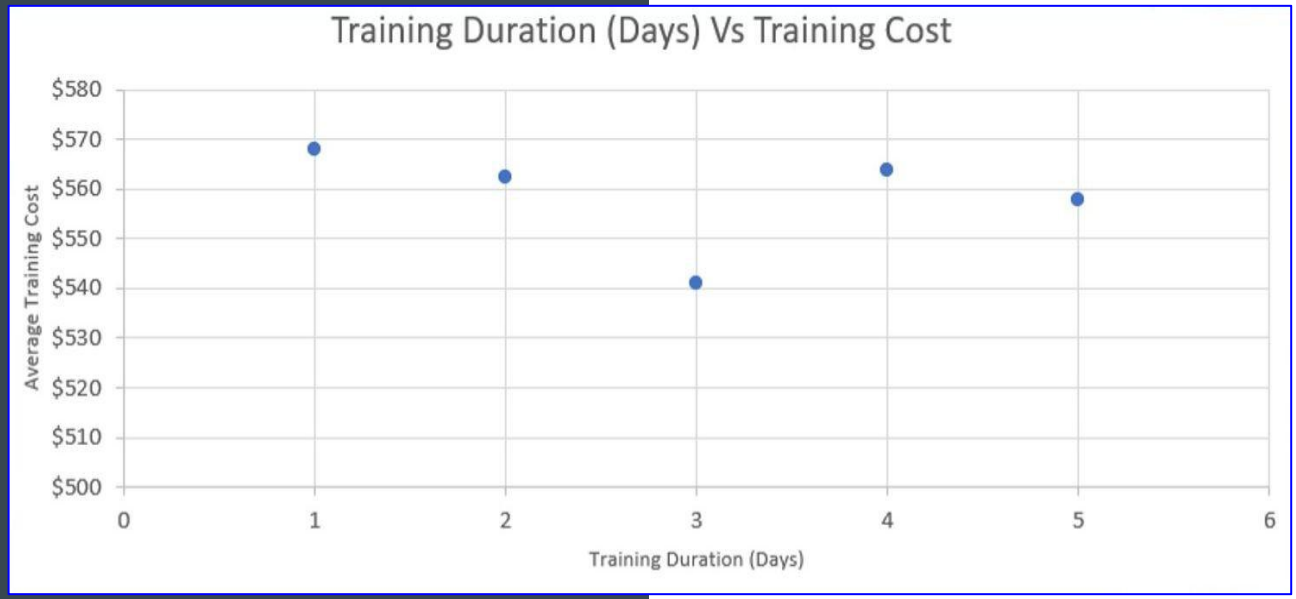
=VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

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

Row Labels	Average of Current Employee Rating
Admin Offices	3.025
Executive Office	2.791666667
IT/IS	2.969767442
Production	2.982178218
Sales	2.909365559
Software Engineering	2.904347826
Grand Total	**2.969**

According to the above table Department with the highest average "Employee Rating" is Admin Offices.

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



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

Count of Employee ID	Column Labels		
Row Labels	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**

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

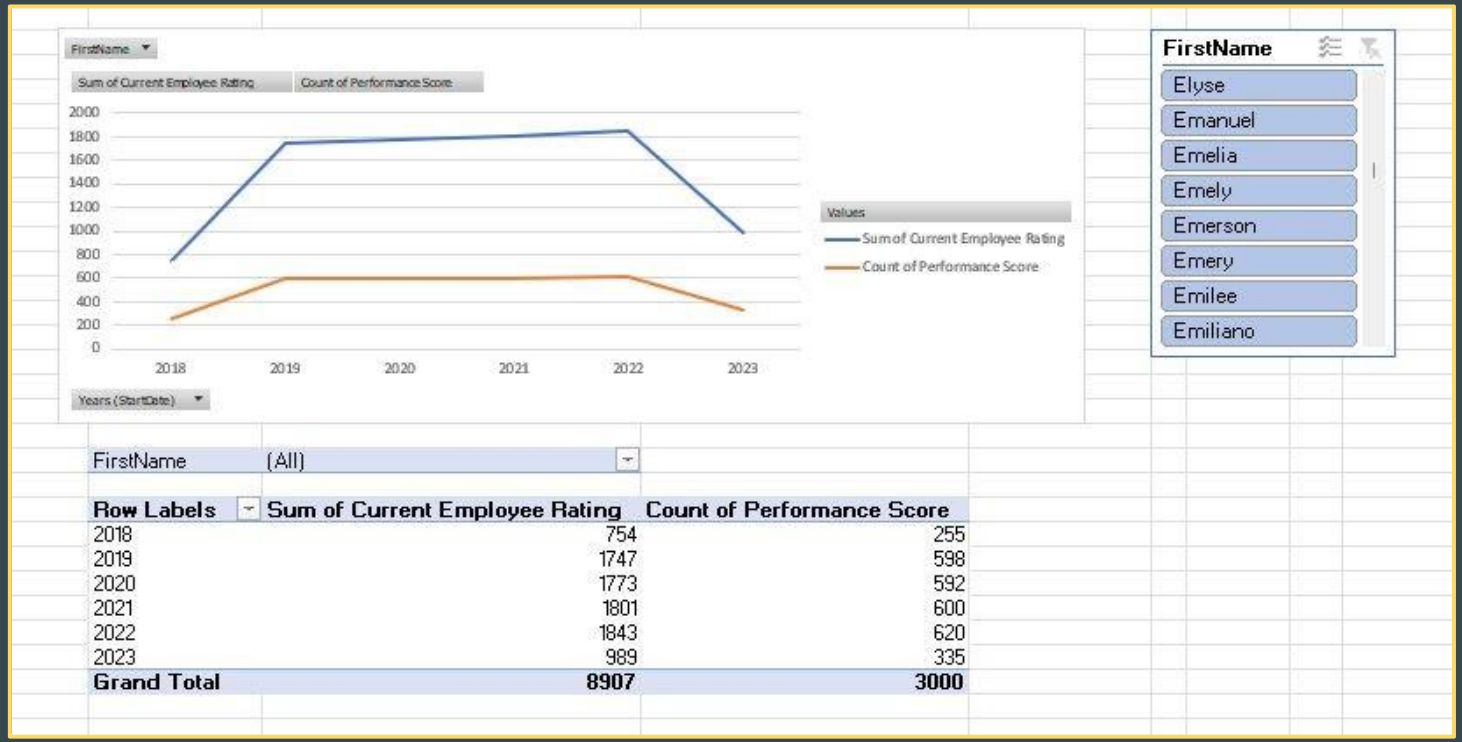
Employee ID	Employee ID
1001	Customer Service
1002	Leadership Development
1003	Technical Skills
1004	Customer Service
1005	Communication Skills
1006	Project Management
1007	Leadership Development
1008	Technical Skills
1009	Customer Service
1010	Communication Skills

=INDEX(array, MATCH(lookup_value, lookup_array, [match_type]), [col_num], [area_num])

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

Job Functions	Count of Performance Score
⊞ BPC	303
⊞ CCDR	300
⊞ EW	302
⊞ MSC	296
⊞ NEL	304
⊞ PL	301
⊞ PYZ	299
⊞ SVG	304
⊞ TNS	297
⊟ WBL	294
Accountant	2
Administration	3
Administrative	7
Administrator	6
Apprentice	1
Billing	2
Civil Hand	1
Clerical	1
Clerk	4
Construction Manager	2
Controller	1
Coordinator	15
Director	4
Drafter	2
Driller	4
Driver	1
Electrician	1
Engineer	35

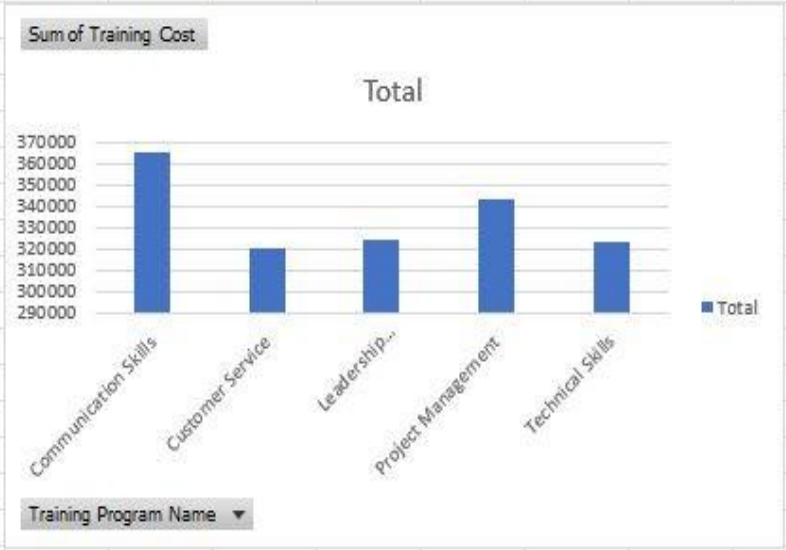
Design a dynamic chart that allows users to select and visualize the performance of any employee over time.



FirstName	(All)	
Row Labels	**Sum of Current Employee Rating**	**Count of Performance Score**
2018	754	255
2019	1747	598
2020	1773	592
2021	1801	600
2022	1843	620
2023	989	335
Grand Total	**8907**	**3000**

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

Row Labels	Sum of Training Cost
Communication Skills	365023.24
Customer Service	320575.04
Leadership Development	323902.03
Project Management	343313.17
Technical Skills	323072.61
Grand Total	**1675886.09**



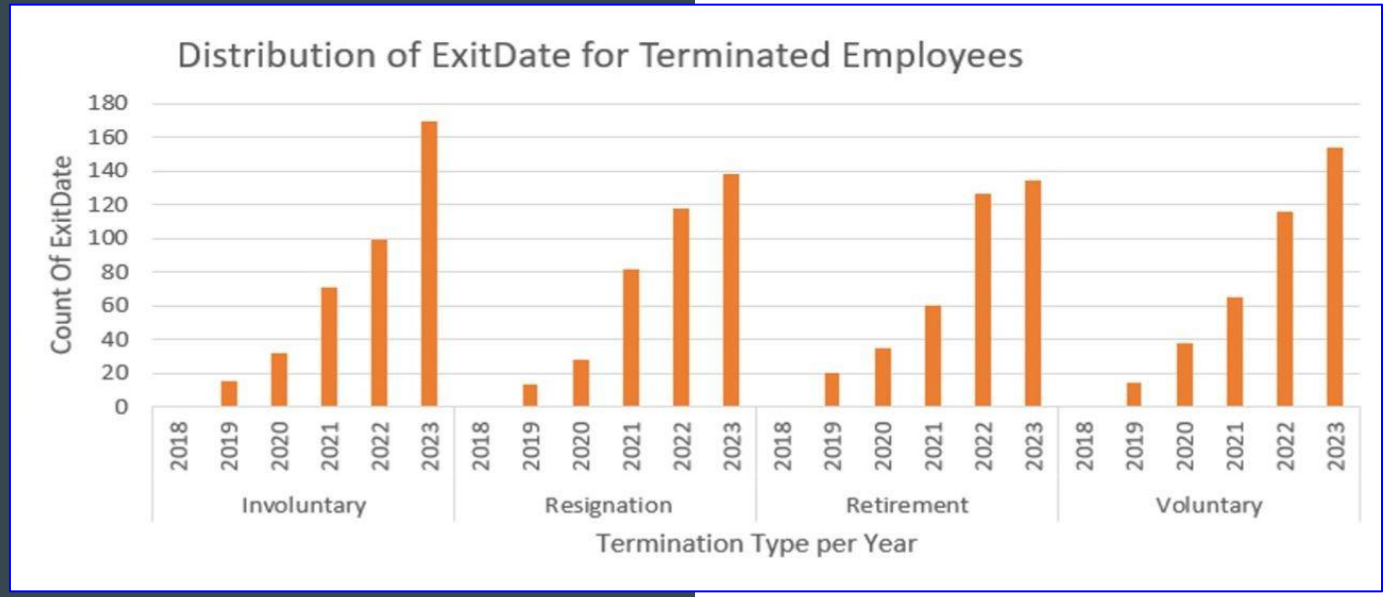
Apply advanced conditional formatting to highlight the top 10% and bottom 10% of employees based on "Current Employee Rating.".

Employee ID	Current Employee Rating	
3427	4	TOP 10%
3428	3	
3429	4	BOTTOM 10%
3430	2	
3431	3	
3432	3	
3433	4	
3434	2	
3435	3	
3436	5	
3437	5	
3438	3	
3439	3	
3440	3	
3441	4	
3442	2	
3443	3	
3444	3	
3445	4	
3446	2	
3447	3	
3448	3	
3449	5	
3450	2	
3451	3	
3452	2	
3453	4	

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

Average of Engagement Score	Column Labels						
Row Labels	2018	2019	2020	2021	2022	2023	Grand Total
Admin Offices	2.875	3.157894737	2.75	3.25	2.545454545	2.7	2.925
Executive Office	4	3.285714286	3.2	3	3.333333333	4	3.375
IT/IS	2.925	3.454545455	2.76	2.908163265	3.261363636	2.463414634	3.025581395
Production	2.835365854	2.950617284	3.020151134	2.788413098	2.925058548	2.852173913	2.906435644
Sales	2.866666667	3.25	2.6875	3.294117647	2.838709677	3.057142857	2.990936556
Software Engineering	3.6	3	3.105263158	3.117647059	2.586206897	2.941176471	2.973913043
Grand Total	**2.898039216**	**3.065217391**	**2.939189189**	**2.888333333**	**2.943548387**	**2.832835821**	**2.939666667**

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



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

Row Labels	Sum of Training Cost
Aaronburgh	633.96
Belltown	555.6
Carlland	104.66
Danielfort	570.13
East Amy	718.16
Fordside	638.69
Griffintown	853.77
Grand Total	**4074.97**
SUM PRODUCT=	**4074.97**

=SUMPRODUCT(array1, [array2], [array3], ...)