PSYUQ

DATA ANALYST INTERNSHIP

HR Data Analysis Assessment:

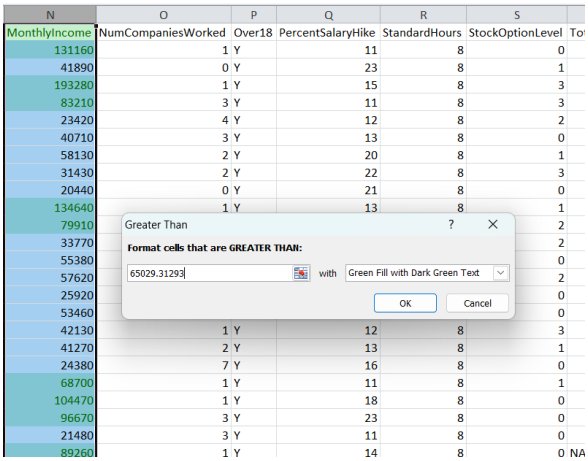
1. **Using Excel, how would you filter the dataset to only show employees aged 30 and above?**
    - Select the Age column
    - Go to data tab
    - Filter the column

	A	B	C	D	E	F	
1	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	Educatio
2	51	No	Travel_Rarely	Sales	6	2	Life Scier
3	31	Yes	Travel_Frequently	Research & Development	10	1	Life Scier
4	32	No	Travel_Frequently	Research & Development	17	4	Other
5	38	No	Non-Travel	Research & Development	2	5	Life Scier
6	32	No	Travel_Rarely	Research & Development	10	1	Medical
7	46	No	Travel_Rarely	Research & Development	8	3	Life Scier
10	31	No	Travel_Rarely	Research & Development	1	3	Life Scier
12	45	No	Travel_Rarely	Research & Development	17	2	Medical
13	36	No	Travel_Rarely	Research & Development	28	1	Life Scier
14	55	No	Travel_Rarely	Research & Development	14	4	Life Scier
15	47	Yes	Non-Travel	Research & Development	1	1	Medical
17	37	No	Travel_Rarely	Research & Development	1	3	Life Scier
19	37	No	Non-Travel	Research & Development	1	3	Medical

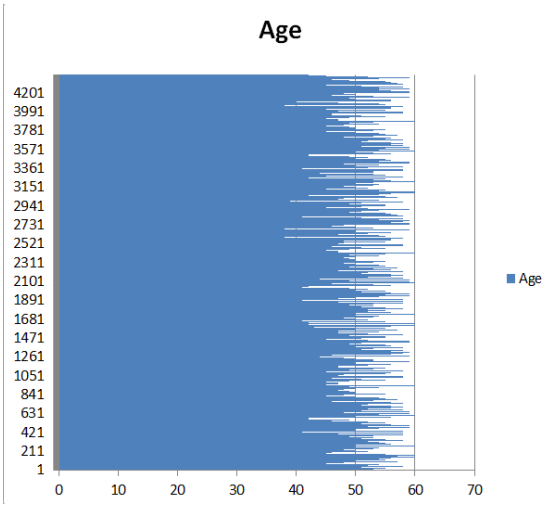
2. **Create a pivot table to summarize the average Monthly Income by Job Role.**

Row Labels	Average of MonthlyIncome
Healthcare Representative	60983.74046
Human Resources	58528.07692
Laboratory Technician	66314.05405
Manager	63395.88235
Manufacturing Director	69183.72414
Research Director	65473.125
Research Scientist	64975.68493
Sales Executive	65186.68712
Sales Representative	65370.96386
(blank)	
Grand Total	**65029.31293**

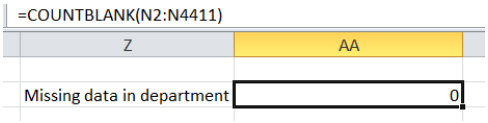
3. **Apply conditional formatting to highlight employees with Monthly Income above the company's average income.**



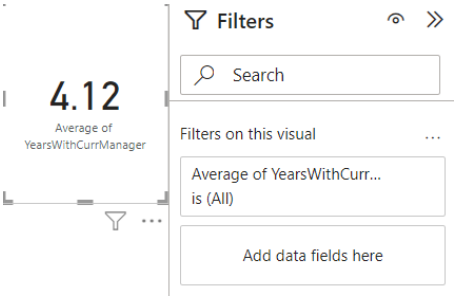
4. **Create a bar chart in Excel to visualize the distribution of employee ages.**



5. **Identify and clean any missing or inconsistent data in the "Department" column.**



6. **Using DAX, create a calculated column that calculates the average years an employee has spent with their current manager.**



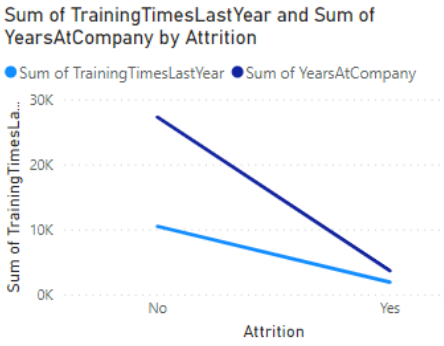
7. **Using Excel, create a pivot table that displays the count of employees in each Marital Status category, segmented by Department.**

Row Labels	Count of EmployeeID
Divorced	**981**
Human Resources	21
Research & Development	621
Sales	339
Married	**2019**
Human Resources	96
Research & Development	1350
Sales	573
Single	**1410**
Human Resources	72
Research & Development	912
Sales	426
(blank)	
Grand Total	**4410**

8. **Apply conditional formatting to highlight employees with both above-average Monthly Income and above-average Job Satisfaction.**

EmployeeID	EnvironmentSatisfaction	JobSatisfaction	WorkLifeBalance	MonthlyIncome
1	3	4	2	131160
2	3	2	4	41890
3	2	2	1	193280
4	4	4	3	83210
5	4	1	3	23420
6	3	2	2	40710
7	1	3	1	58130
8	1	2	3	31430
9	2	4	3	20440
10	2	1	3	134640
11	3	4	3	79910
12	NA	4	3	33770
13	4	1	3	55380
14	1	2	2	57620
15	4	4	2	25920
16	3	4	4	53460
17	4	3	4	42130
18	1	4	3	41270
19	2	2	2	24380
20	1	1	3	68700
21	3	2	1	104470
22	1	2	2	96670
23	3	3	2	21480
24	2	3	3	89260
25	2	4	2	65130
26	2	4	3	67990
27	1	1	3	162910
28	4	4	3	27050
29	4	3	1	103330
30	4	4	3	44480

9. **In Power BI, create a line chart that visualizes the trend of Employee Attrition over the years.**



10. **Describe how you would create a star schema for this dataset, explaining the benefits of doing so.**

A star schema for an employee dataset might look like this:

Fact Table:

- **EmployeePerformance Columns:** EmployeeID (Foreign Key to EmployeeDimension), ManagerID (Foreign Key to ManagerDimension), Date, MonthlyIncome, JobSatisfaction, etc.

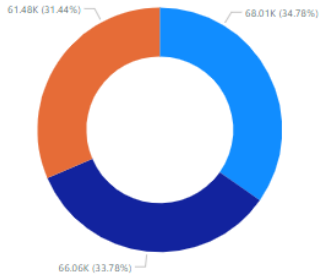
Dimension Tables:

- **EmployeeDimension Columns:** EmployeeID (Primary Key), EmployeeName, DepartmentID, HireDate, etc.
- **ManagerDimension Columns:** ManagerID (Primary Key), ManagerName, HireDate, etc.
- **DateDimension Columns:** DateID (Primary Key), Date, Day, Month, Quarter, Year, etc.
- **DepartmentDimension Columns:** DepartmentID (Primary Key), DepartmentName, Location, etc.

Benefits of Star Schema:

- Simplicity and ease of understanding.
- Improved query performance due to efficient aggregations.
- Scalability with flexibility for adding or modifying dimensions.
- Easier navigation for end-users and enhanced tool compatibility.
- Separation of concerns between measures and descriptive attributes.
- Enhanced data integrity through foreign key relationships.

11. Using DAX, calculate the rolling 3-month average of Monthly Income for each employee.



12. Create a hierarchy in Power BI that allows users to drill down from Department to Job Role to further narrow their analysis.

13. How can you set up parameterized queries in Power BI to allow users to filter data based on the Distance from Home column?

- Create a parameter in the Manage Parameters window.
- In Power Query Editor, apply a filter to the "Distance from Home" column using the created parameter.
- Load and apply changes to the data.
- Create a visual in your Power BI report.
- Add a slicer for the parameter in the Slicer visualization.
- Users can interactively use the slicer to set the distance filter, dynamically updating the data in their reports.

14. In Excel, calculate the total Monthly Income for each Department, considering only the employees with a Job Level greater than or equal to 3.

Row Labels	Count of MonthlyIncome
1	**1629**
Human Resources	57
Research & Development	1098
Sales	474
2	**1602**
Human Resources	72
Research & Development	1032
Sales	498
3	**654**
Human Resources	30
Research & Development	435
Sales	189
4	**318**
Human Resources	12
Research & Development	192
Sales	114
5	**207**
Human Resources	18
Research & Development	126
Sales	63
(blank)	
(blank)	
Grand Total	**4410**

15. Explain how to perform a What-If analysis in Excel to understand the impact of a 10% increase in Percent Salary Hike on Monthly Income.

- Input your data with columns like "Percent Salary Hike" and "Monthly Income."
- Add an input cell for the original Percent Salary Hike value.
- Create a formula to calculate the new Monthly Income based on the percent salary hike
- Set up a table or section for the What-If analysis with different values for Percent Salary Hike.
- Copy the formula across rows to calculate Monthly Income for each What-If scenario.
- Optionally, create a chart to visualize the impact.
- Optionally, use the Scenario Manager for managing and comparing different scenarios.

16. Verify if the data adheres to a predefined schema. What actions would you take if you find inconsistencies?

- Review the predefined schema.
- Inspect the data for deviations.
- Use data profiling tools to analyze characteristics.
- Check column data types, data ranges, and constraints.
- Address missing values and standardize formats.
- Implement data transformations as needed.
- Document and communicate any inconsistencies found.
- Coordinate with data stewards or domain experts.
- Iterate through verification and correction processes.
- Establish ongoing data quality monitoring.
- Communicate findings and corrections to stakeholders.