# Data Wrangling

# Data Dynamos
# (Our Critical Discussion)

For HR Data Analysis

# Data Wrangling

❖ **Our Role in Data Wrangling:**
- Our primary responsibility in the HR Analysis project was focused on Data Wrangling.
- Data Wrangling involves cleaning, organizing, and structuring raw data to make it usable for analysis and insights.
- We combined and prepared datasets.

❖ **Why Data Wrangling is Critical:**
- Data Wrangling is a fundamental process because raw data often comes in complex and unstructured formats.
- By cleaning and organizing the data, we ensure that the project has a solid foundation for accurate analysis.
- We ensured that relevant data is accessible in one dataset.

❖ **Main and Supporting Datasets:**
- The core of our work involved processing a main dataset with past HR Analysis.
- Supporting datasets included:
  - ☒ The dataset appears to have a wealth of HR-related columns. Key fields include:
    - **Demographics: Gender, Age, Ethnicity, MaritalStatus**
    - **Job Details: Department, DistanceFromHome, BusinessTravel, HireDate**
    - **Performance and Tenure: PerformanceRating, YearsAtCompany, YearsInMostRecentRole, YearsSinceLastPromotion, YearsWithCurrManager**
    - **Compensation: Salary, StockOptionLevel**
    - **Other: OverTime, Attrition**

❖ **Merging Data for Better Insights:**
- We used merge operations to integrate datasets.
- This process made it easier to analyze key factors affecting each element.

❖ **Working with Nested Data:**
- In case of nested data, we will use data explosion techniques to convert these lists into rows, making each element analyzable individually.
- This allowed us to compute metrics.

❖ **Summary of Our Discussion & Contributions:**

☒ **Key Contributions:**

- Cleaned and organized data from multiple HR datasets.
- Merged essential data points.
- Handled complex data structures to make the data ready for analysis.

❖ **Summary of Our Discussion & Contributions:**

☒ **Key Contributions:**

**Data Dynamos**