St. Francis Institute of Technology, Mumbai-400 103

**Department Of Information Technology**

A.Y. 2024-2025

Class: TE-ITA/B, Semester: VI

Subject: **Business Intelligence Lab**

**Business Intelligence Mini Project**

1. **Aim:** Develop a Business IntelligenceMini Project for a particular case study
2. **Objectives:** After study of this experiment, the students will be able to develop mini project
3. **Outcomes:**

**CO6:** Apply BI to solve practical problems: Analyze the problem domain, use the data collected in enterprise apply the appropriate data mining technique, interpret and visualize the results and provide decision support

1. **Prerequisite:** Study of DM&BI Tools.
2. **Requirements:** Personal Computer, Windows XP operating system/Windows 7, Internet Connection, Microsoft Word, WEKA tool, Orange tool, BI Tool.
3. **Theory: Nil**
4. **Laboratory Exercise:** Each group select one case study for this. A BI report (Initial pages shared separately) must be prepared outlining the following steps:
5. Write the problem statement for your case study
6. Draw star schema and snowflake schema
7. Give dataset details, identifying which data mining task is needed
8. Download and use a standard data mining dataset available for the problem. Some links for data mining datasets are WEKA site, UCI Machine Learning Repository, KDD site, KDD Cup etc.
9. Implement the data mining algorithm using Weka and Orange
10. Interpret and visualize the results using BI tool like Qlikview & Tableau
11. After interpretation clearly provide the BI decision that is to be taken

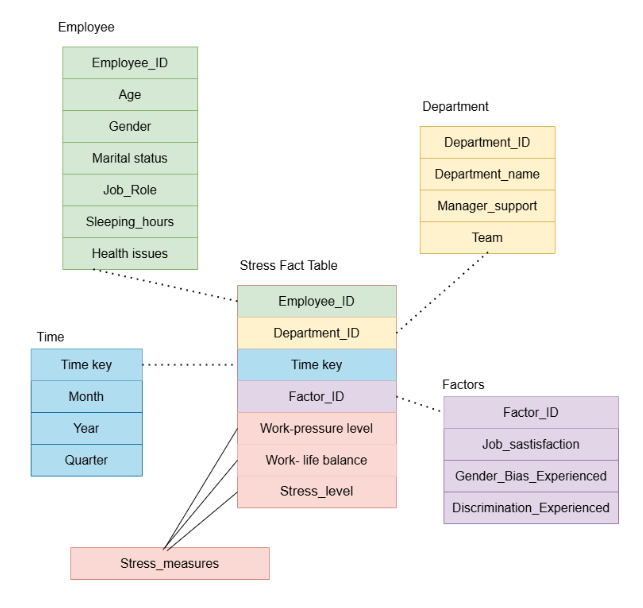
1. **Post-Experiment Exercise:**
2. **Conclusion:**
   * Summary of mini project
3. **Reference:** Business Intelligence: Data Mining and Optimization for Decision Making by Carlo Vercellis, Wiley India Publication

**Problem statement:**

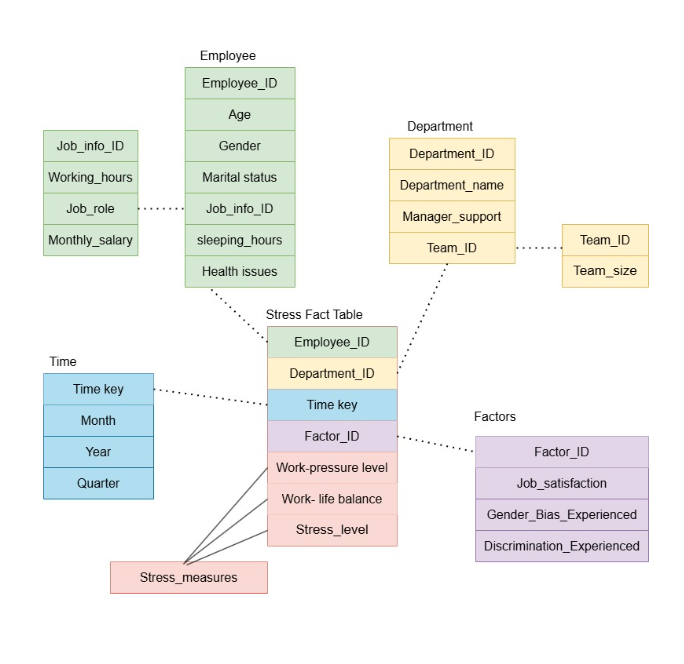
Corporate stress is an escalating concern in today’s high-pressure work environments, where relentless deadlines and overwhelming expectations leave employees grappling with exhaustion and burnout. This chronic stress doesn't just erode individual well-being—it saps productivity, dampens creativity, and undermines organizational success. Yet, many organizations remain ill-equipped to recognize and address these invisible struggles. The availability of a corporate stress dataset offers a glimmer of hope, a chance to uncover the hidden patterns behind workplace stress and empower organizations to foster healthier, more compassionate work environments.

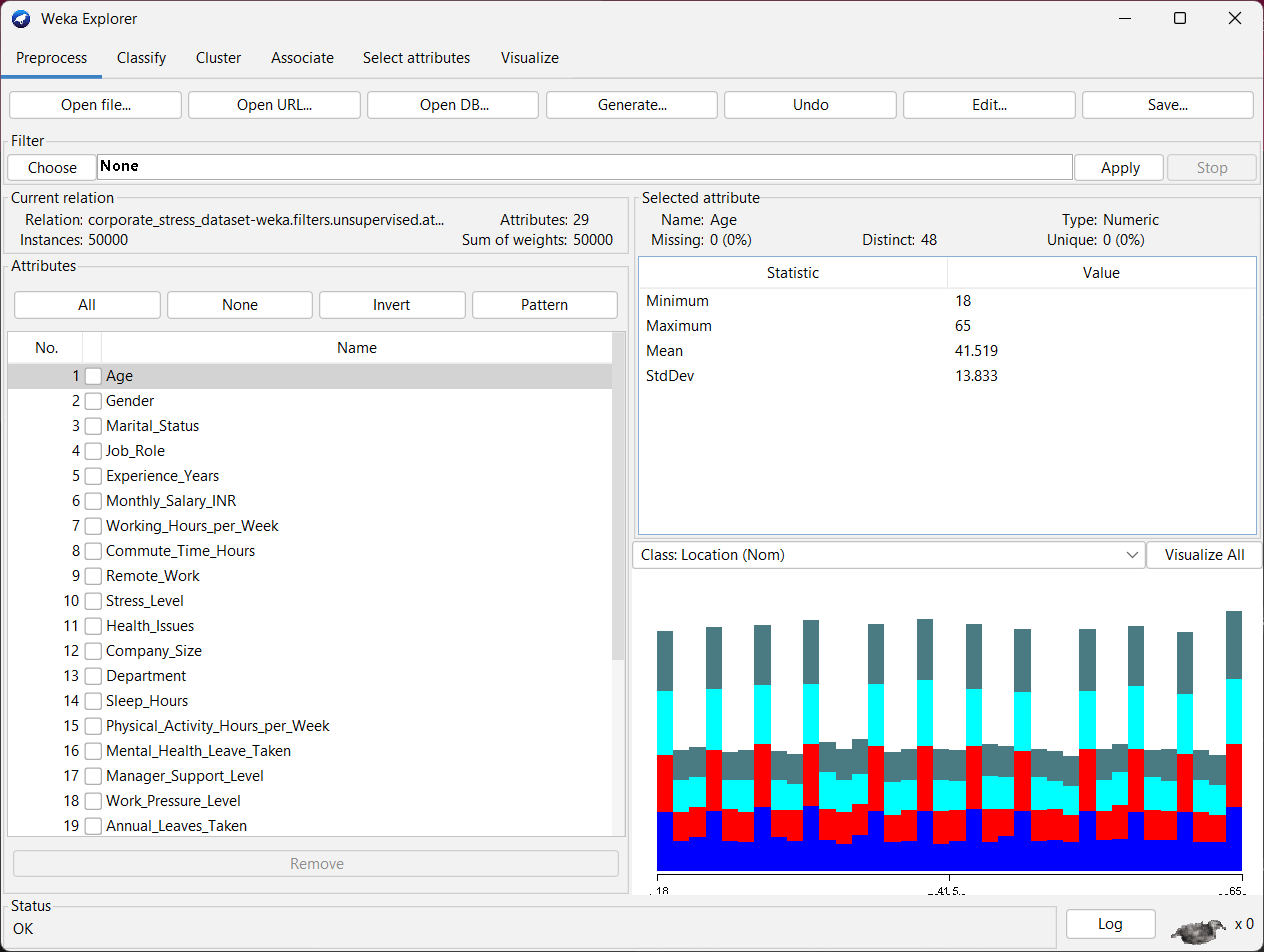
By harnessing this dataset with advanced analytics and machine learning, we can illuminate the connections between stress and factors like excessive work hours, insufficient managerial support, and inequitable policies. These insights are not just numbers—they are a call to action. They can guide organizations in crafting tailored strategies, from personalized stress management programs to equitable workload distribution. The ultimate goal is to replace burnout with balance, turning workplaces into spaces where individuals thrive and collective success is inevitable.

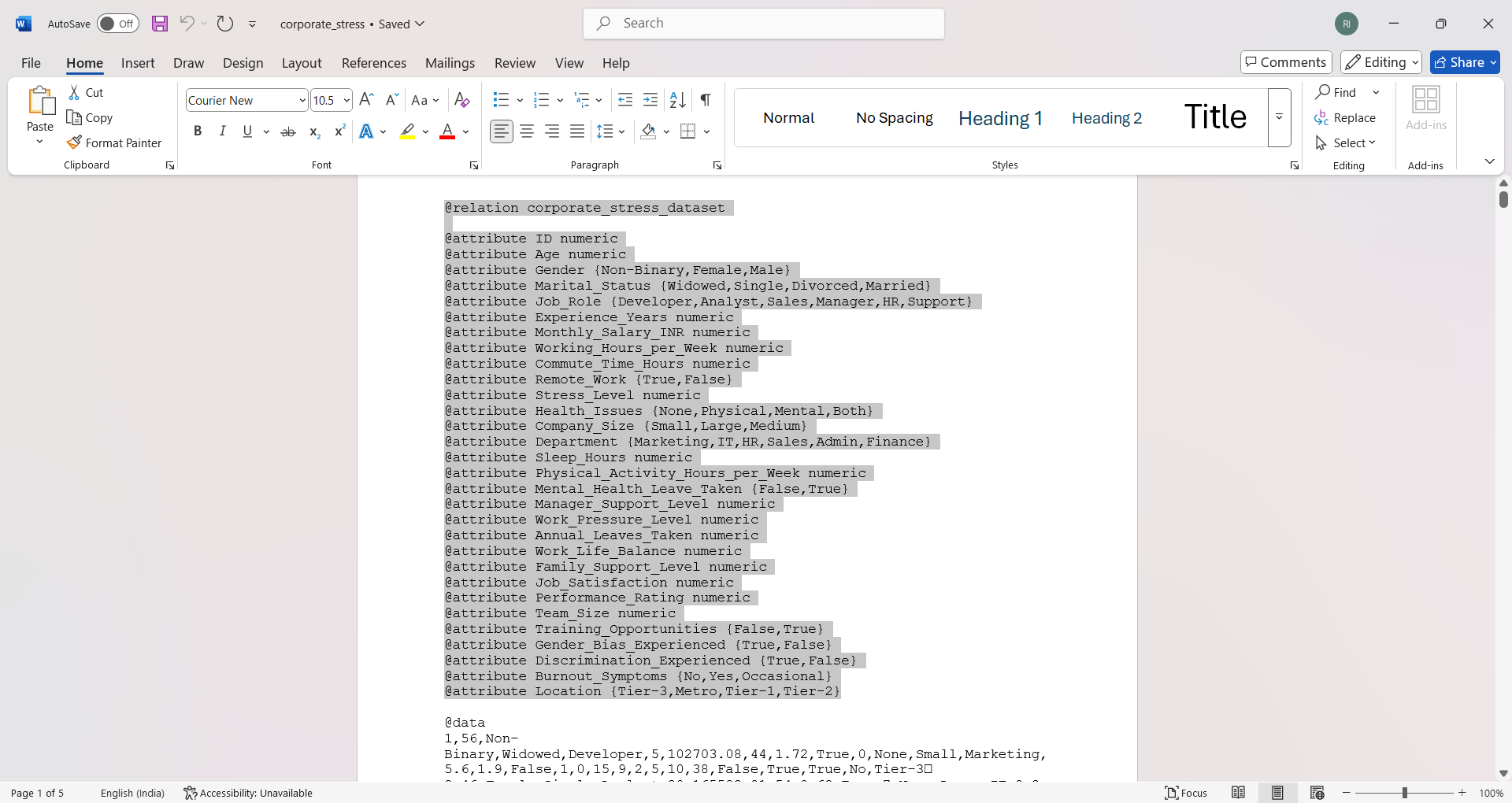
**Star-Schema:**



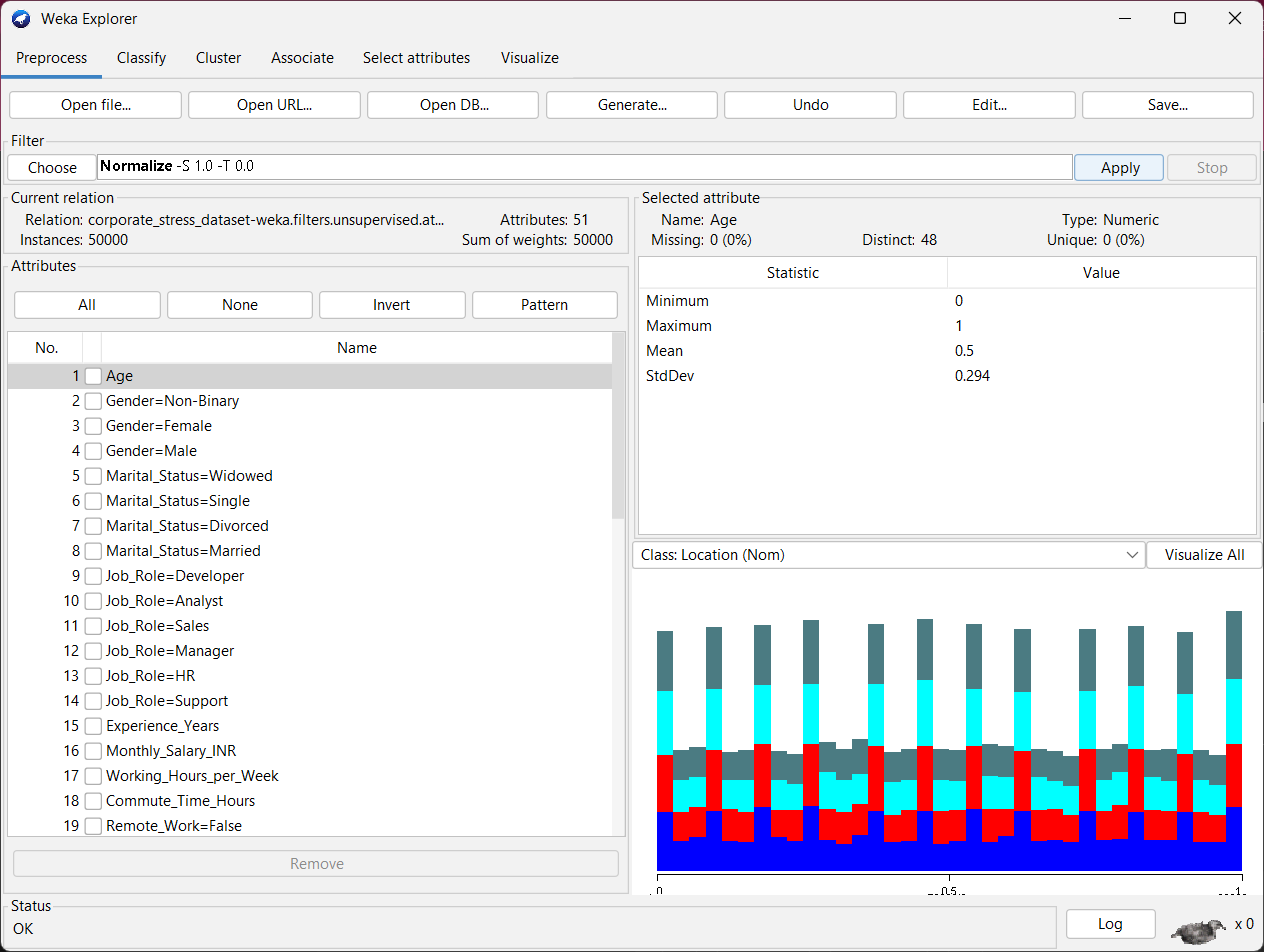
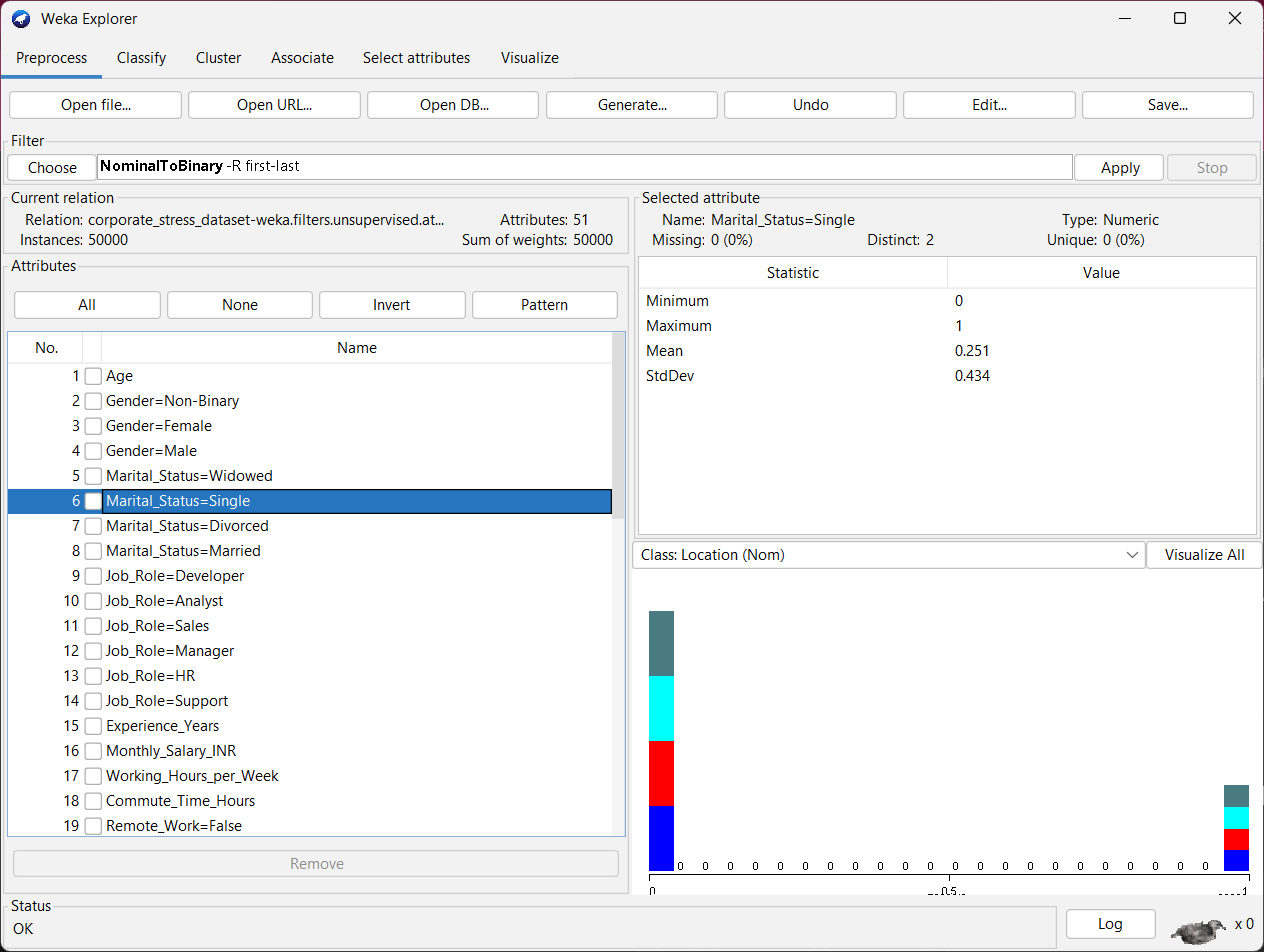
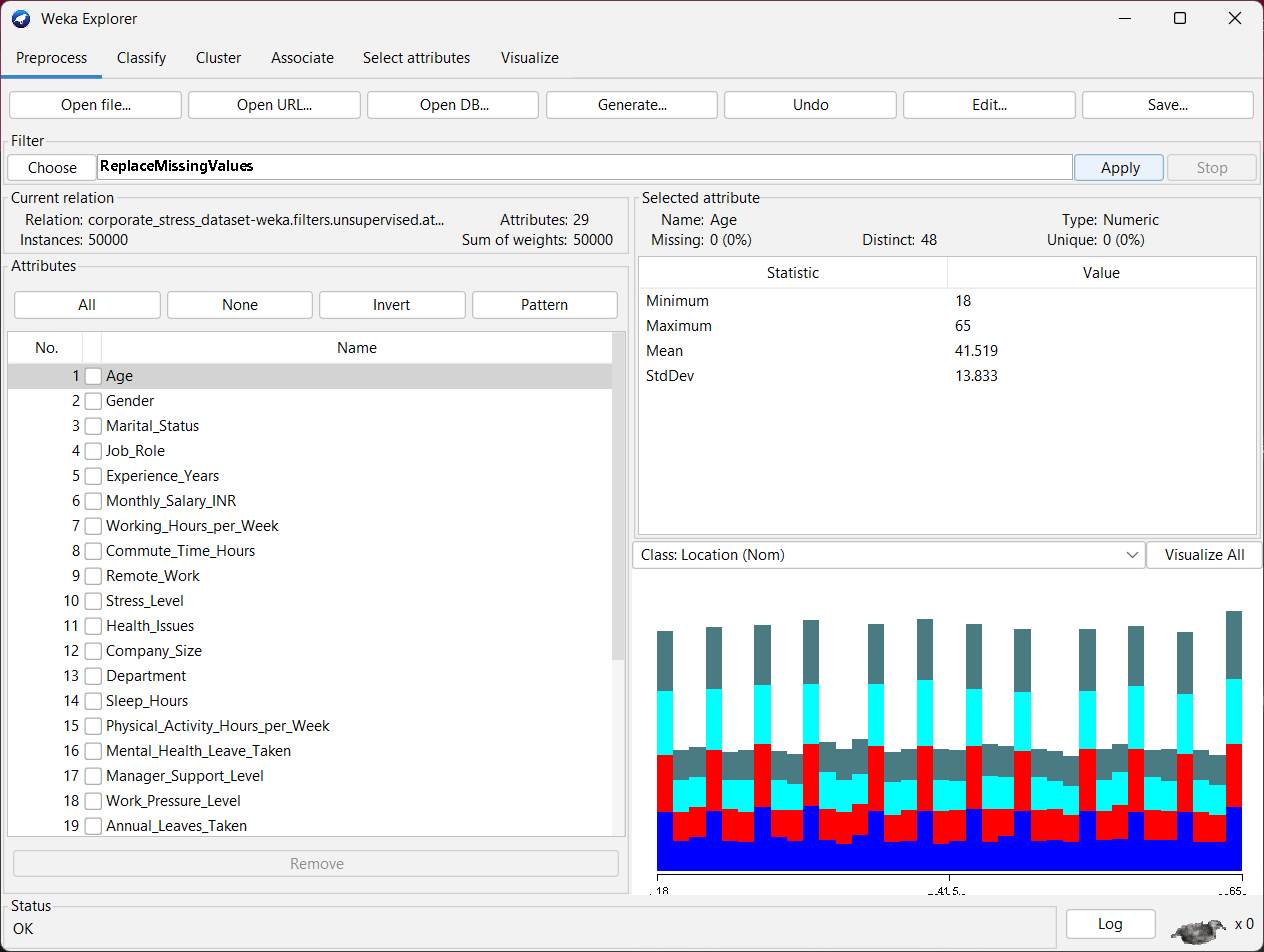
**Snowflake:**



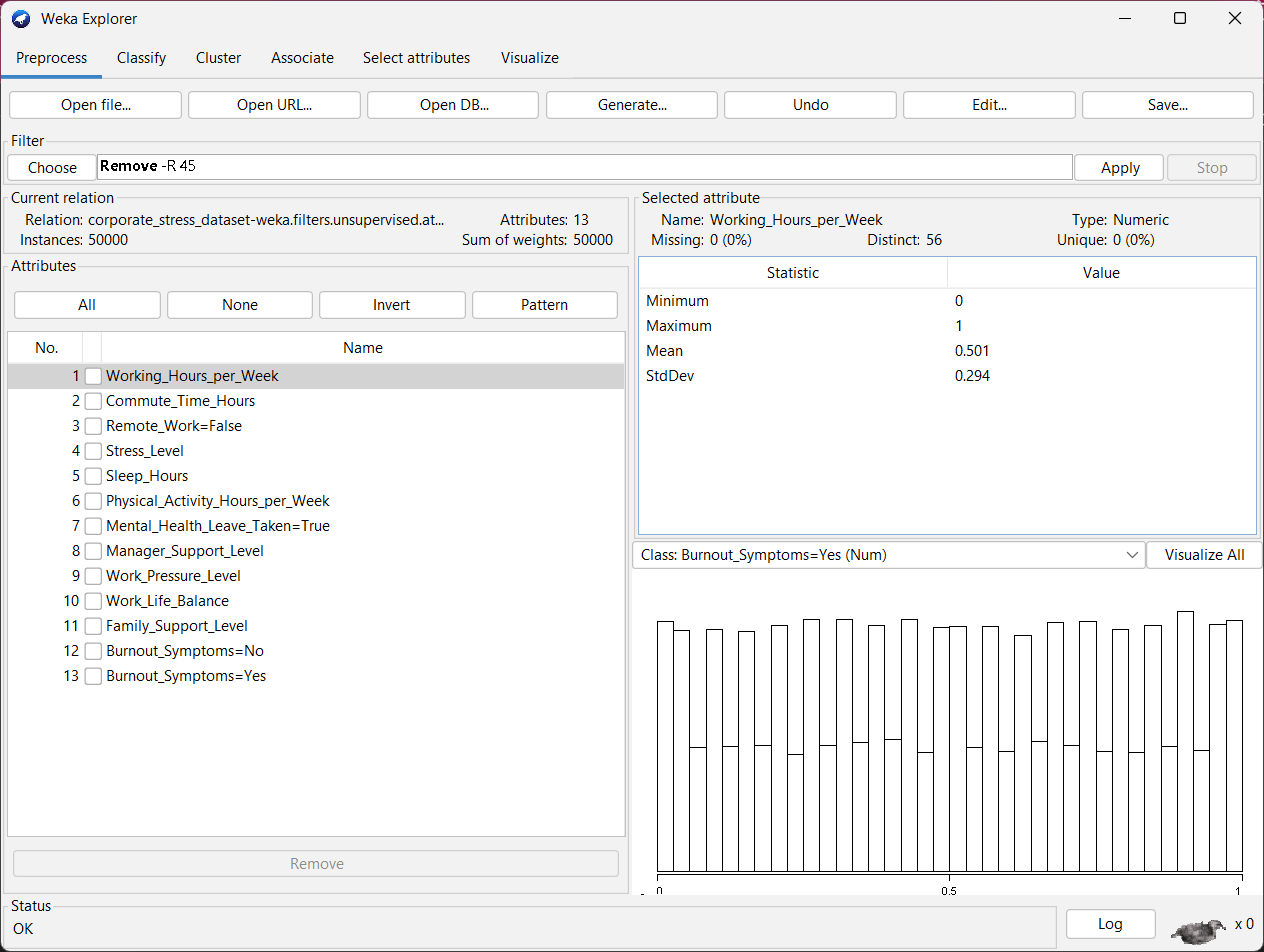
Loading dataset in WEKA  


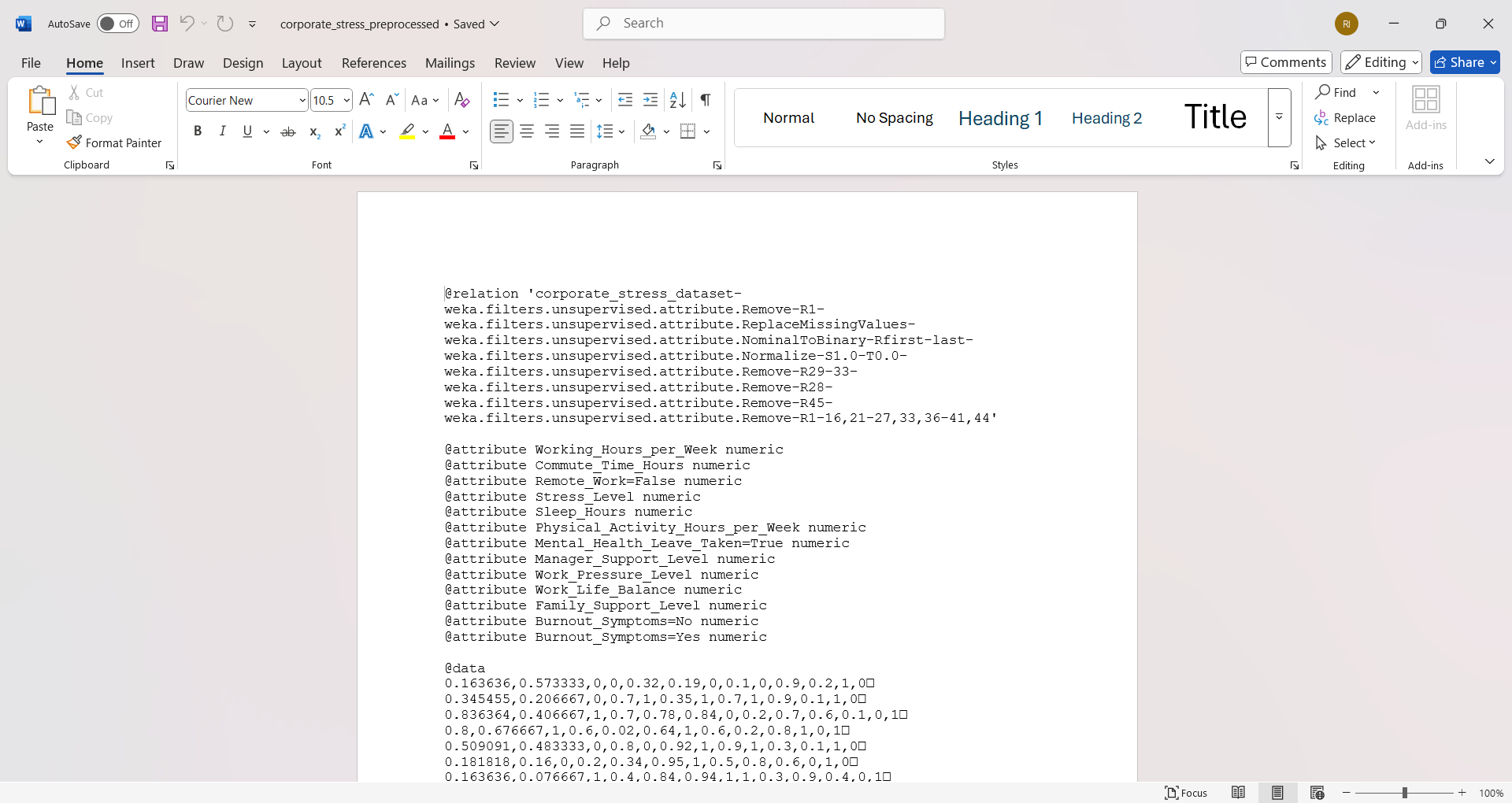


Preprocessing dataset

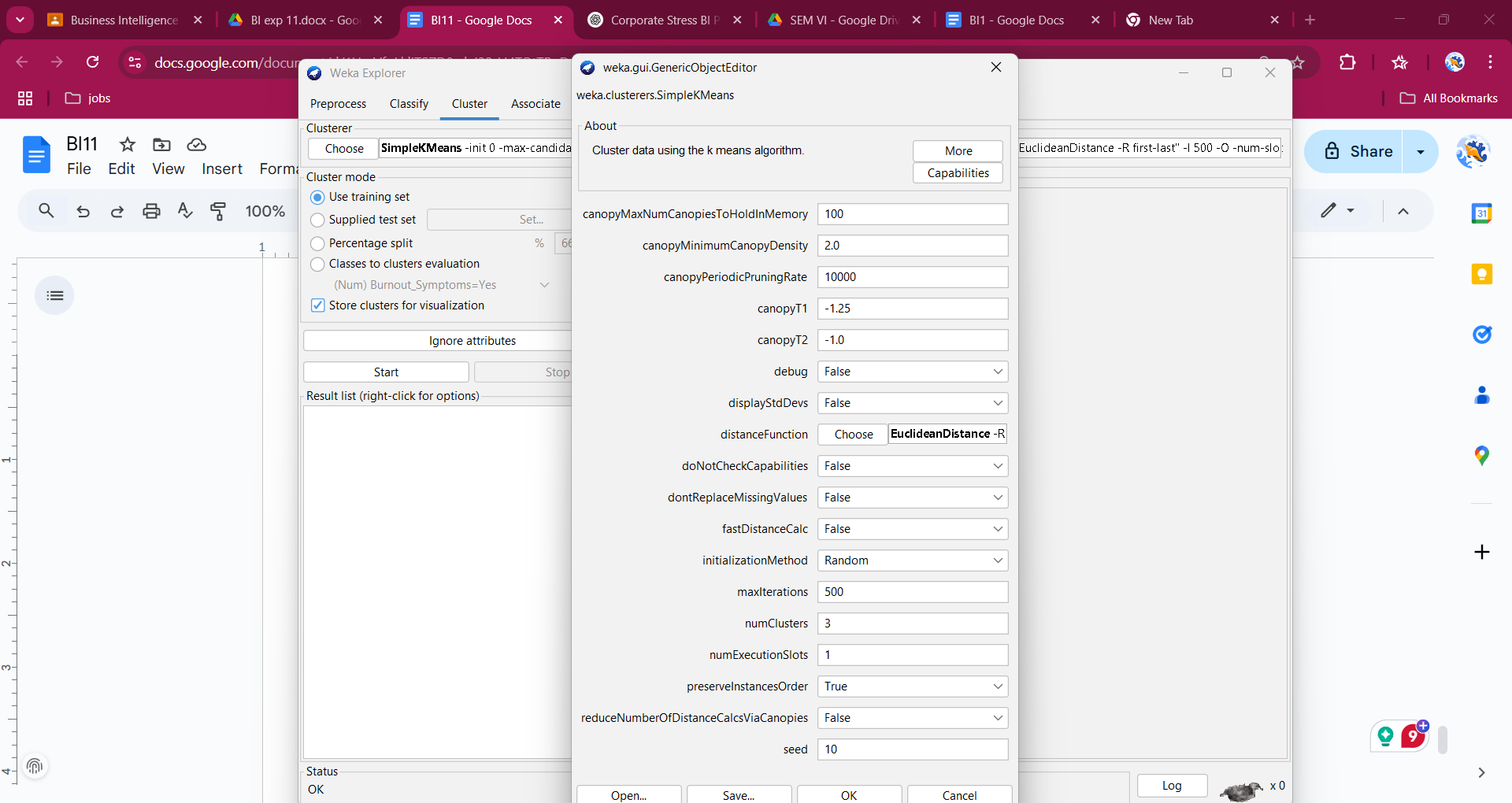


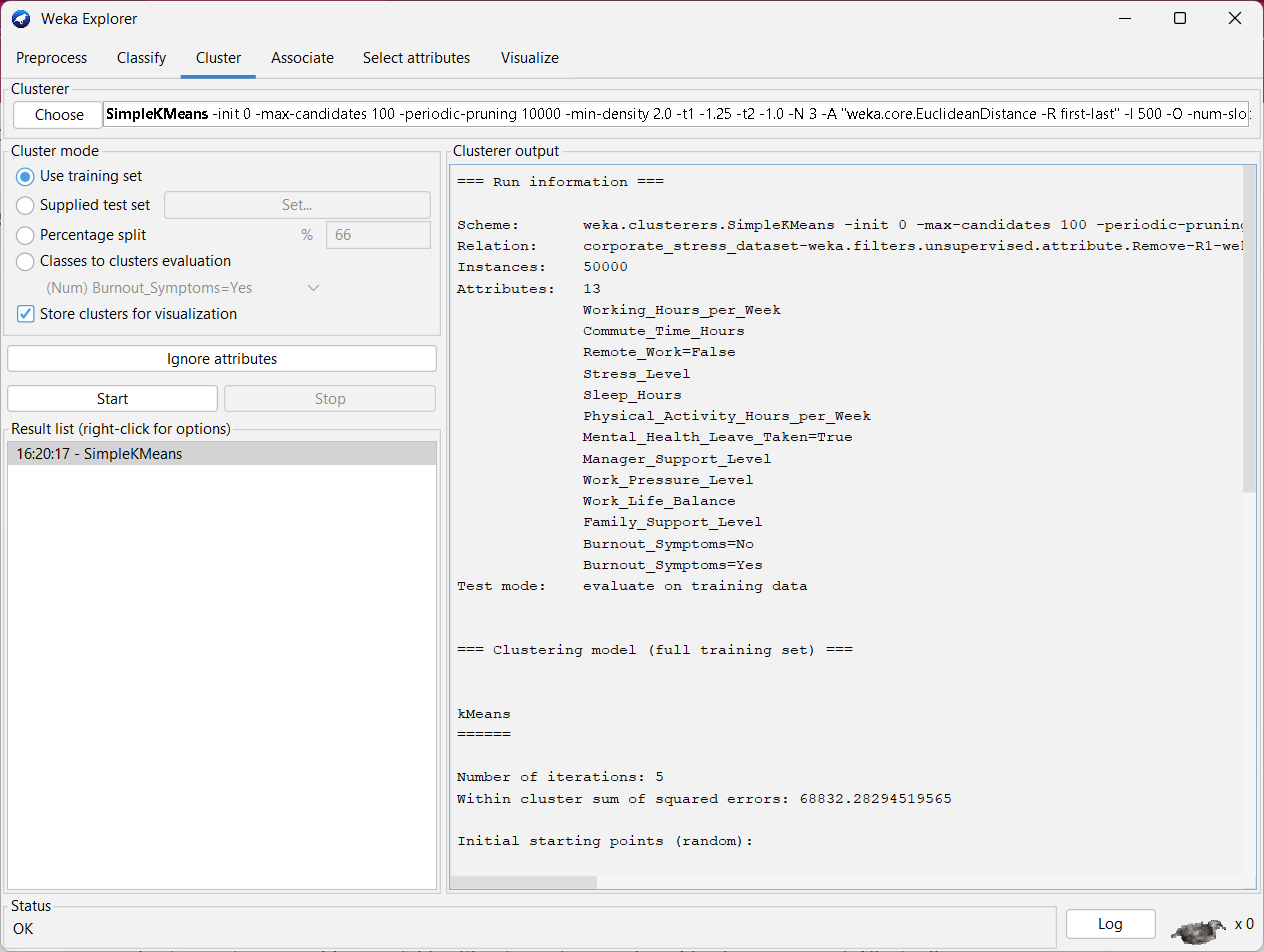
dataset ready for clustering

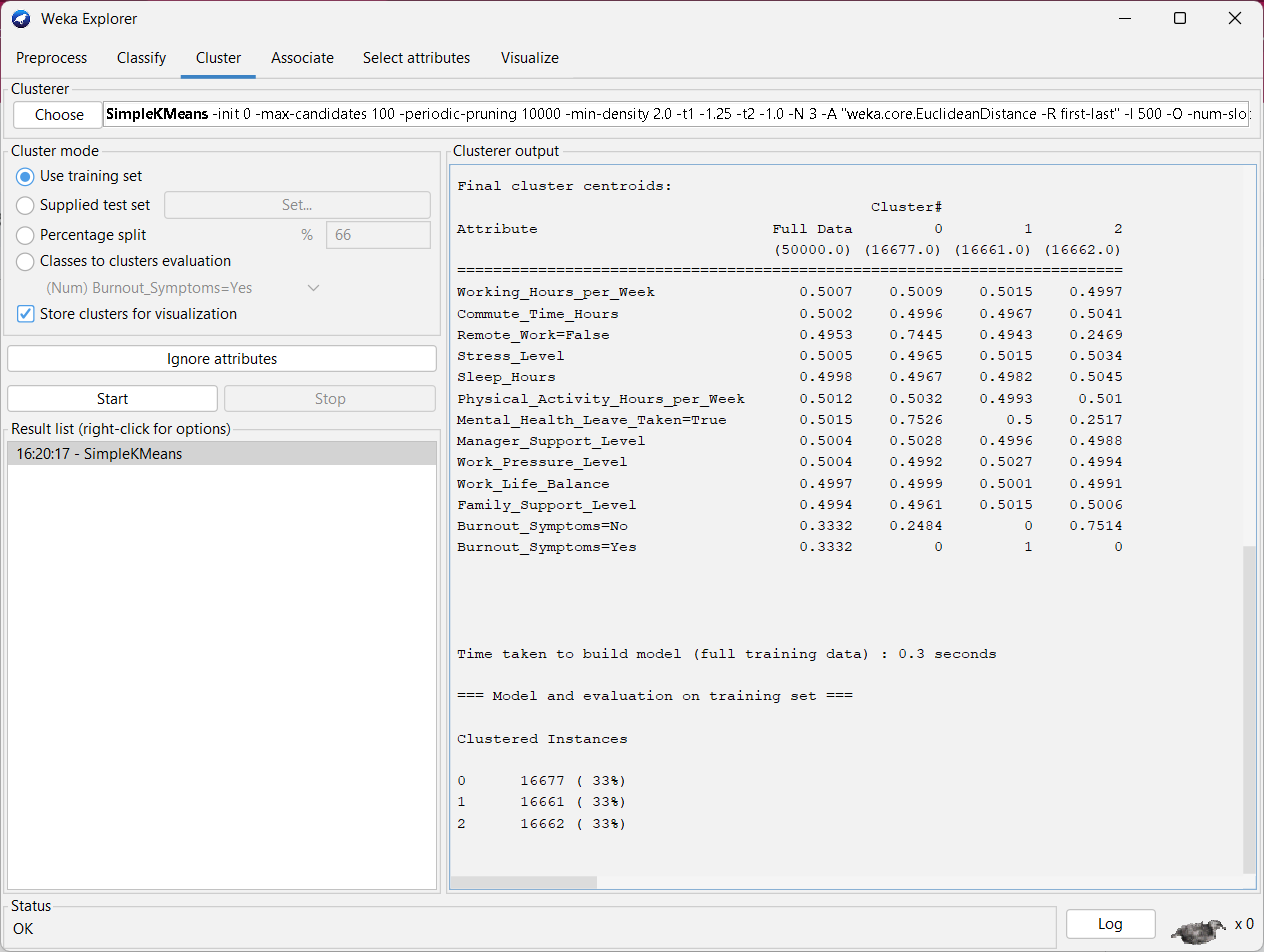




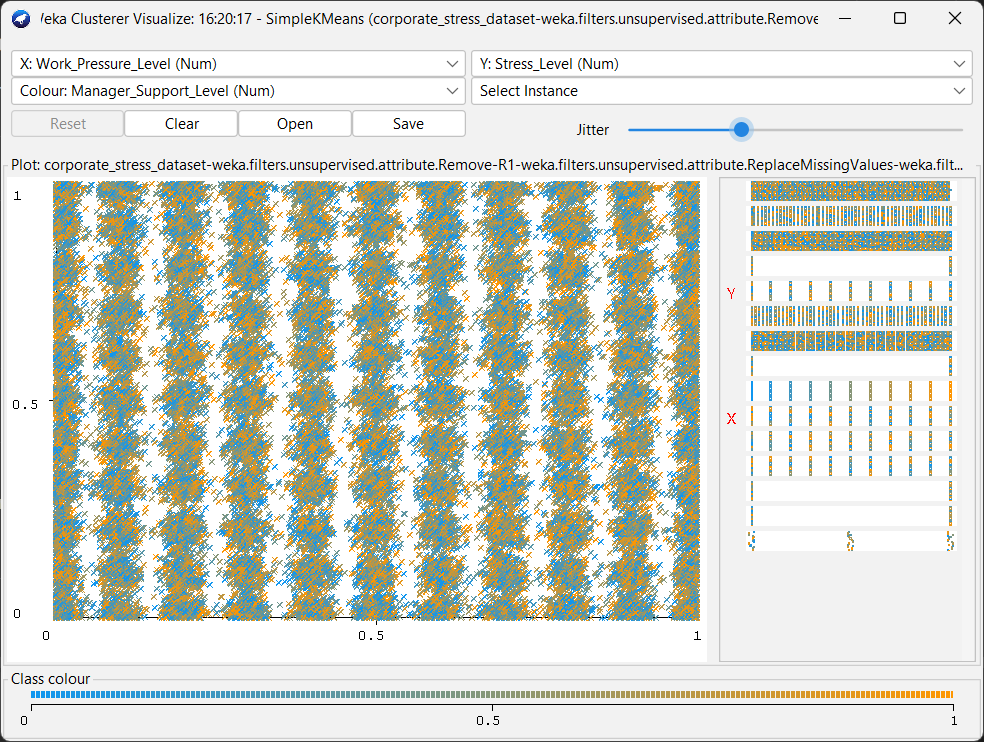
clustering using K-means





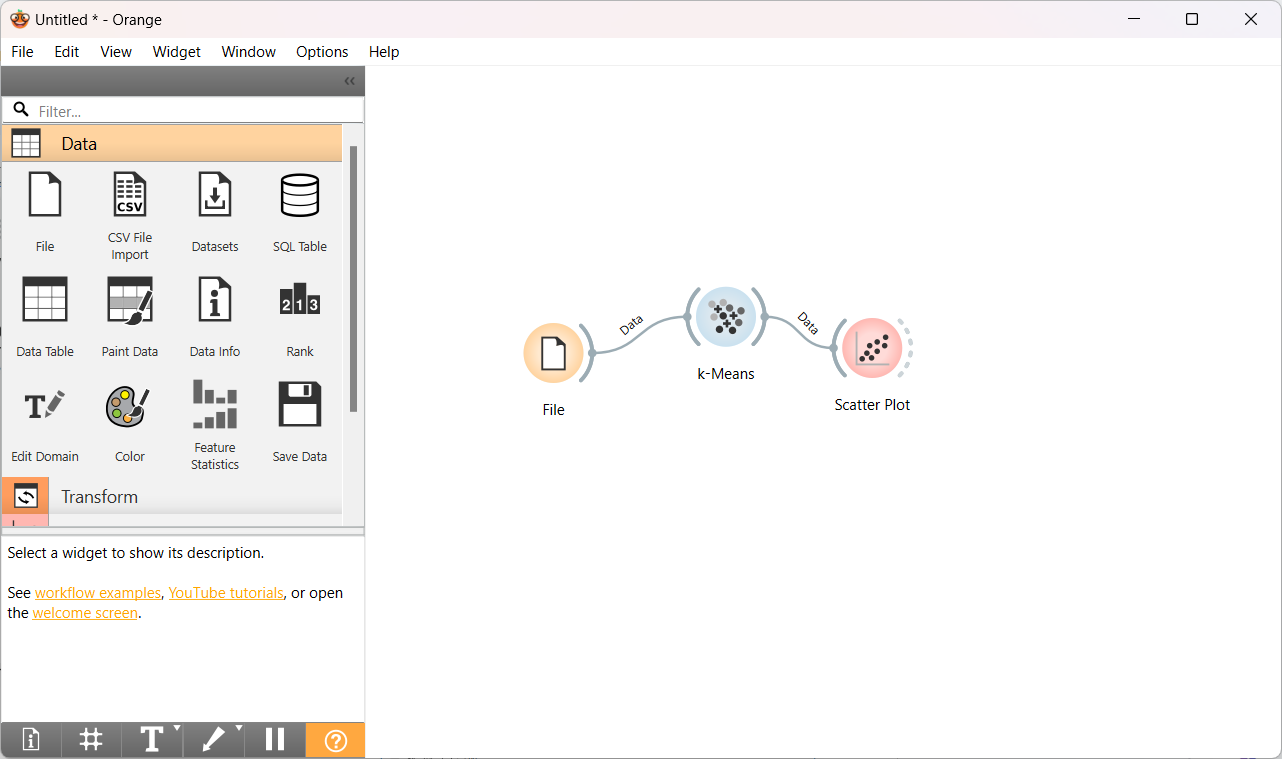


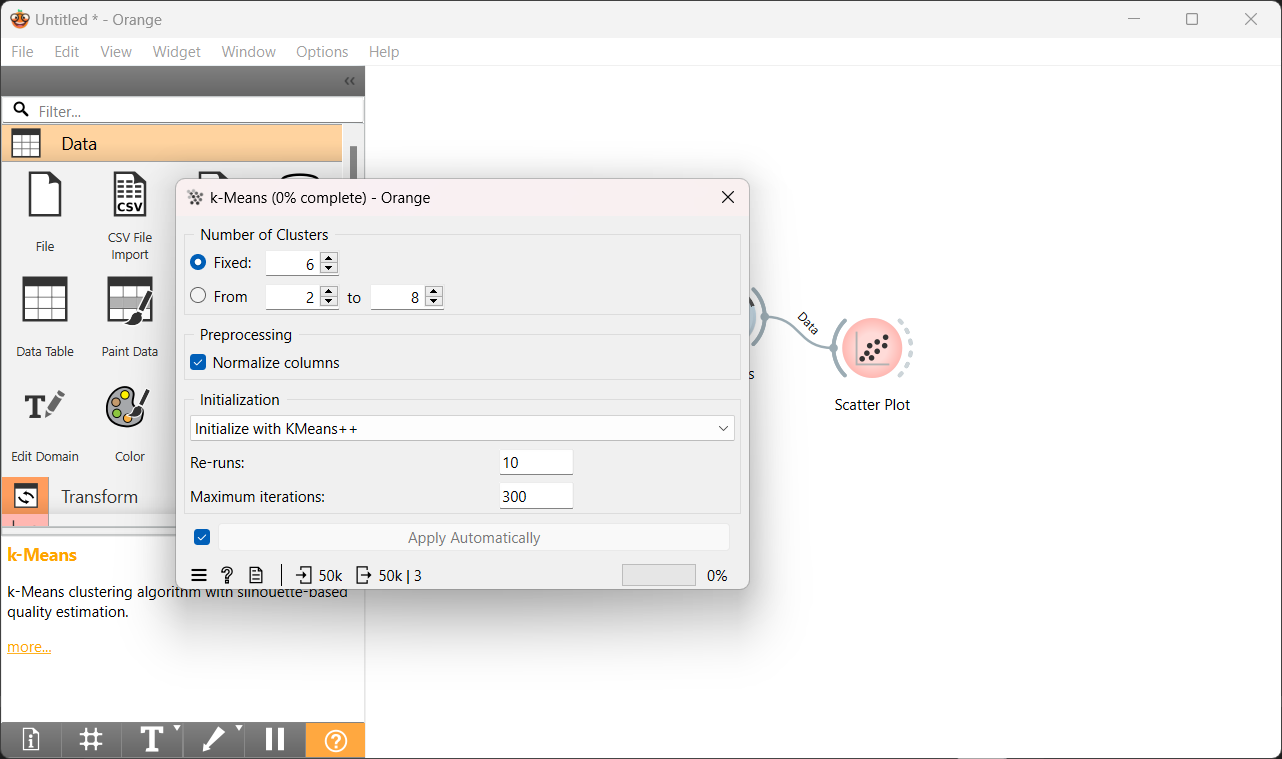
The K-Means clustering applied to the corporate stress dataset revealed three distinct employee groups based on behavioral and stress-related attributes. Cluster 1 emerged as the most critical group, with 100% of its members showing burnout symptoms and relatively high stress levels, indicating a strong need for mental health interventions and workload reassessment. Cluster 0 included employees who had mostly taken mental health leave and were not working remotely, suggesting they are at risk and may require ongoing support. In contrast, Cluster 2 represented a relatively resilient segment, with the majority showing no burnout symptoms, working remotely, and maintaining better sleep patterns. These insights can guide targeted wellness programs, improved managerial practices, and policy adjustments to foster a healthier, more supportive work environment.

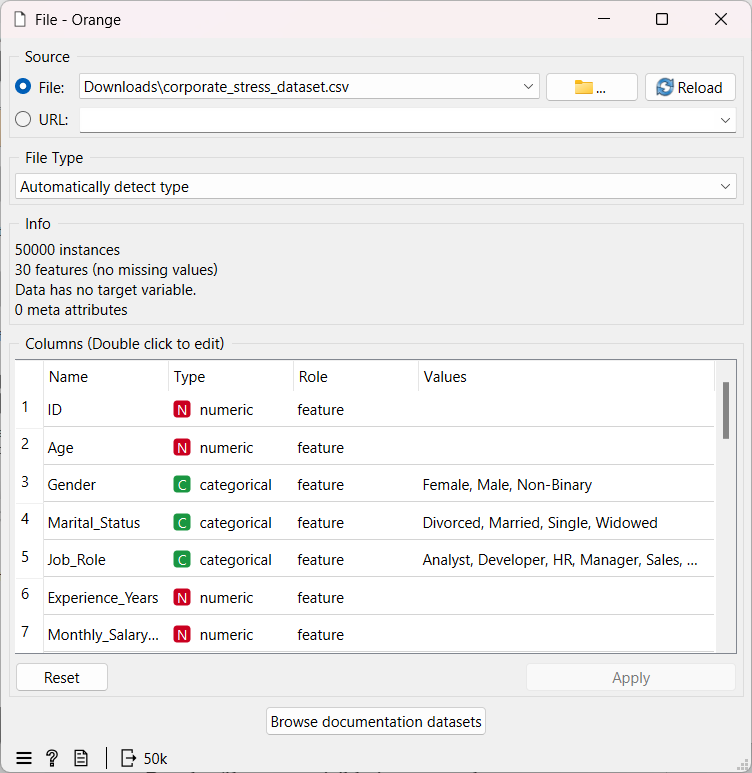


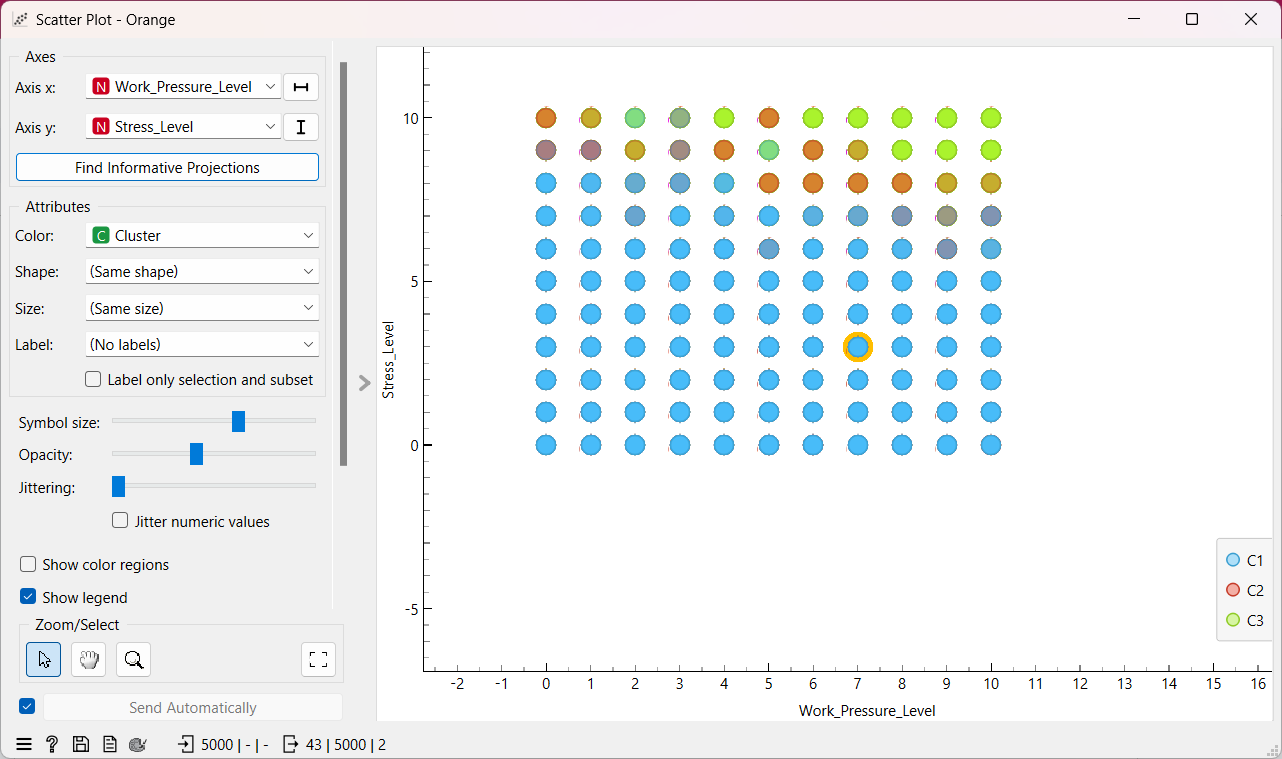
The cluster visualization illustrates the relationship between Work\_Pressure\_Level and Stress\_Level, with colors representing variations in Manager\_Support\_Level. The dense, striped pattern suggests limited variation in work pressure, while stress levels vary more broadly. Color patterns hint that higher manager support (bluer shades) may correspond to lower stress, whereas lower support (yellow/orange shades) aligns with higher stress. However, the overlap between clusters implies that these attributes alone may not clearly separate groups, and exploring other attribute combinations could reveal more distinct clustering.

Performing K-means in Orange



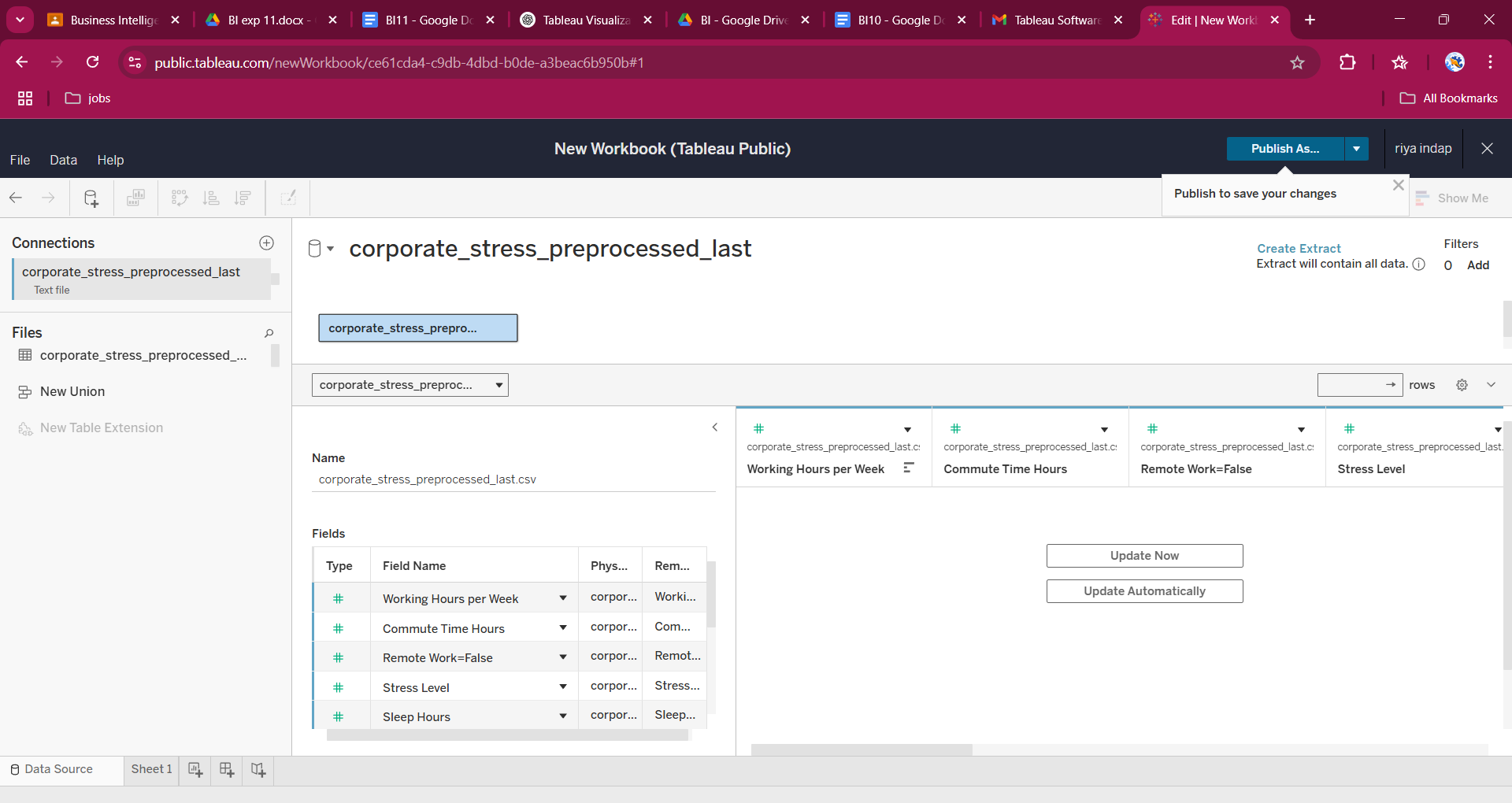
Loading dataset in Orange: Performing K-means algorithm for cluster size=6

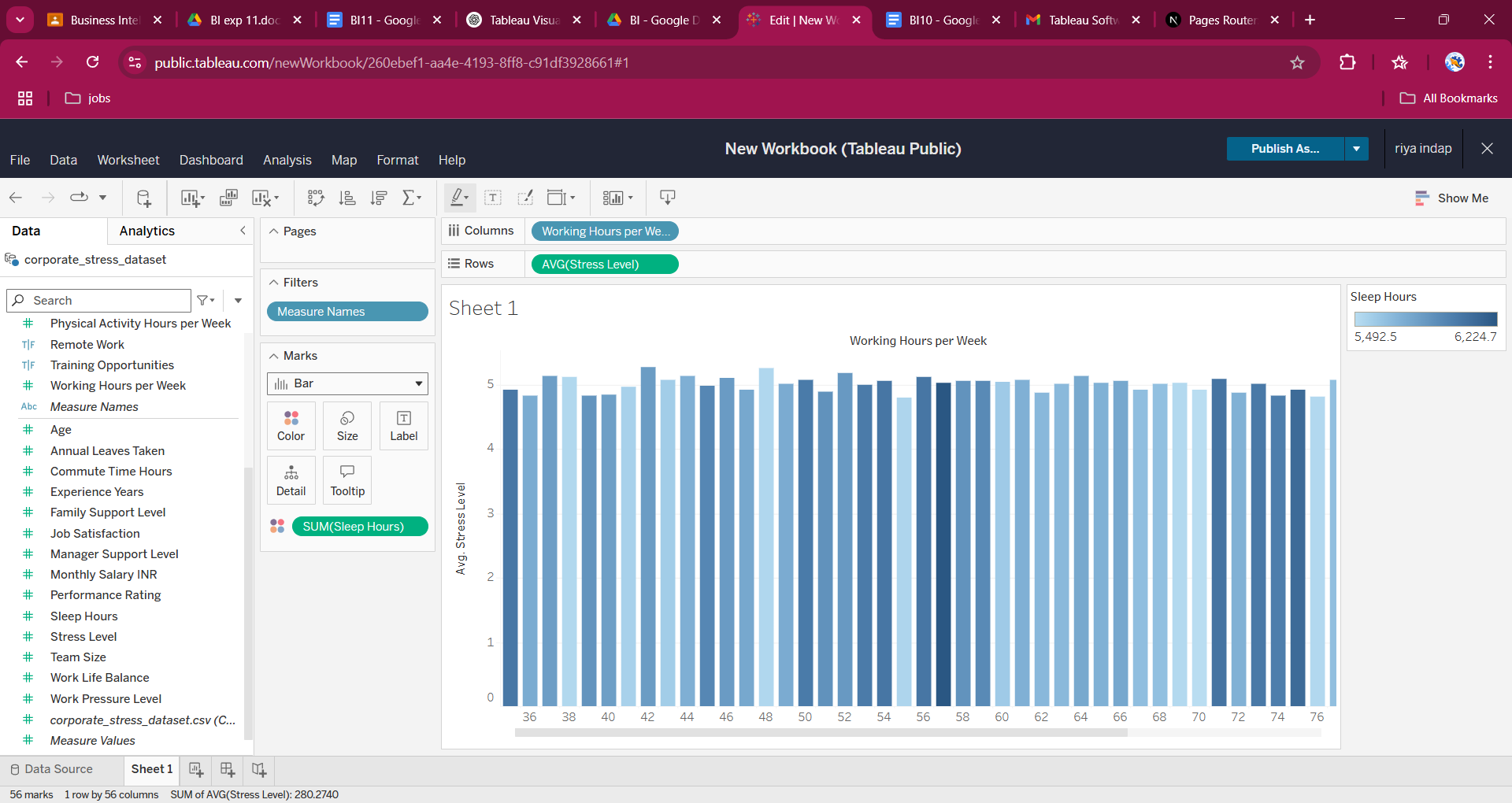




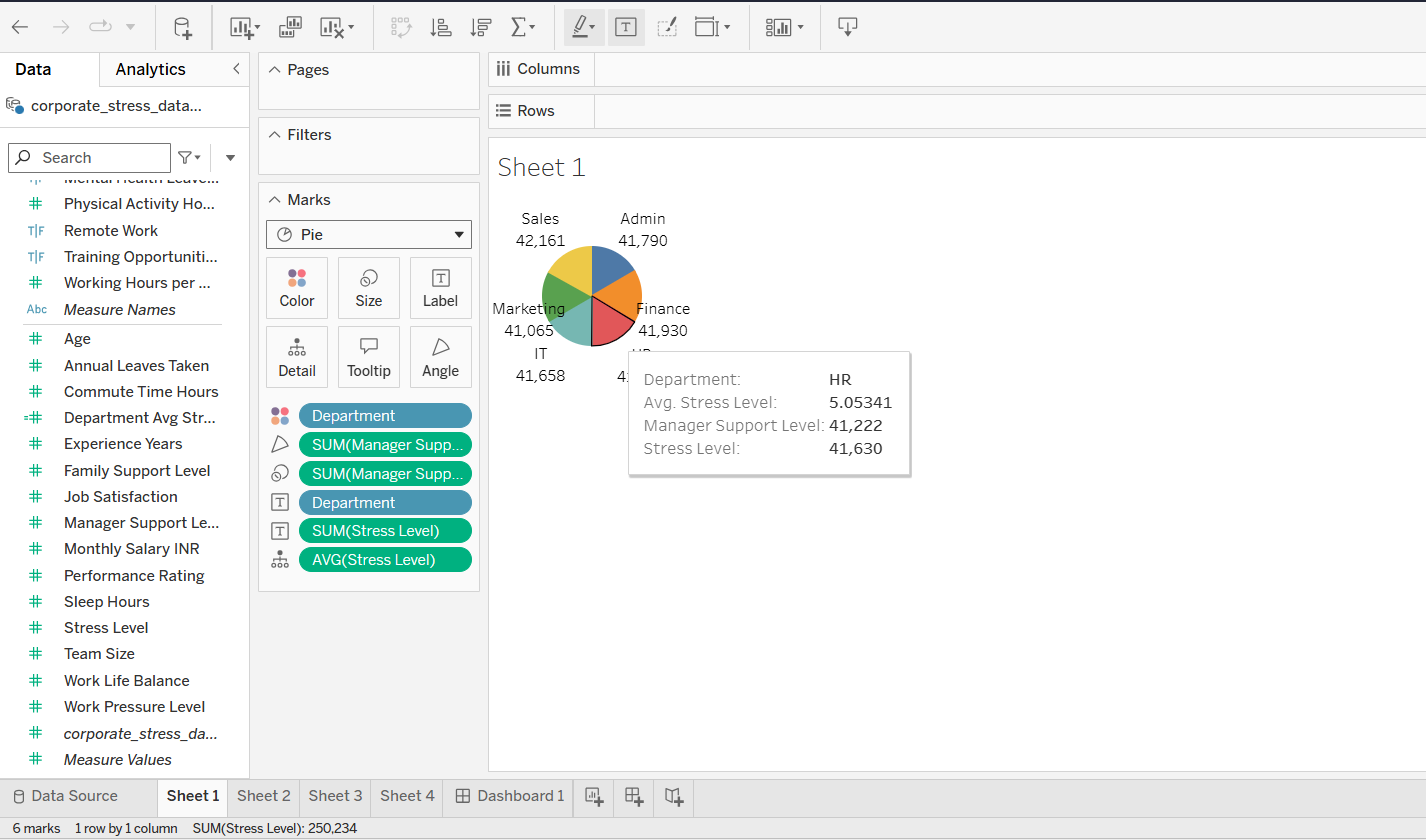
The scatter plot reveals three distinct clusters based on the stress and work pressure levels. Cluster C1 (blue) consists of individuals with moderate stress and pressure, dominating the dataset. Cluster C2 (red) indicates a group with high work pressure and elevated stress, potentially suggesting burnout. Cluster C3 (green) reflects individuals experiencing high stress levels with varying degrees of work pressure, implying other sources of stress. While some overlap exists, the clusters show identifiable behavioral patterns.

Loading dataset in Tableau:

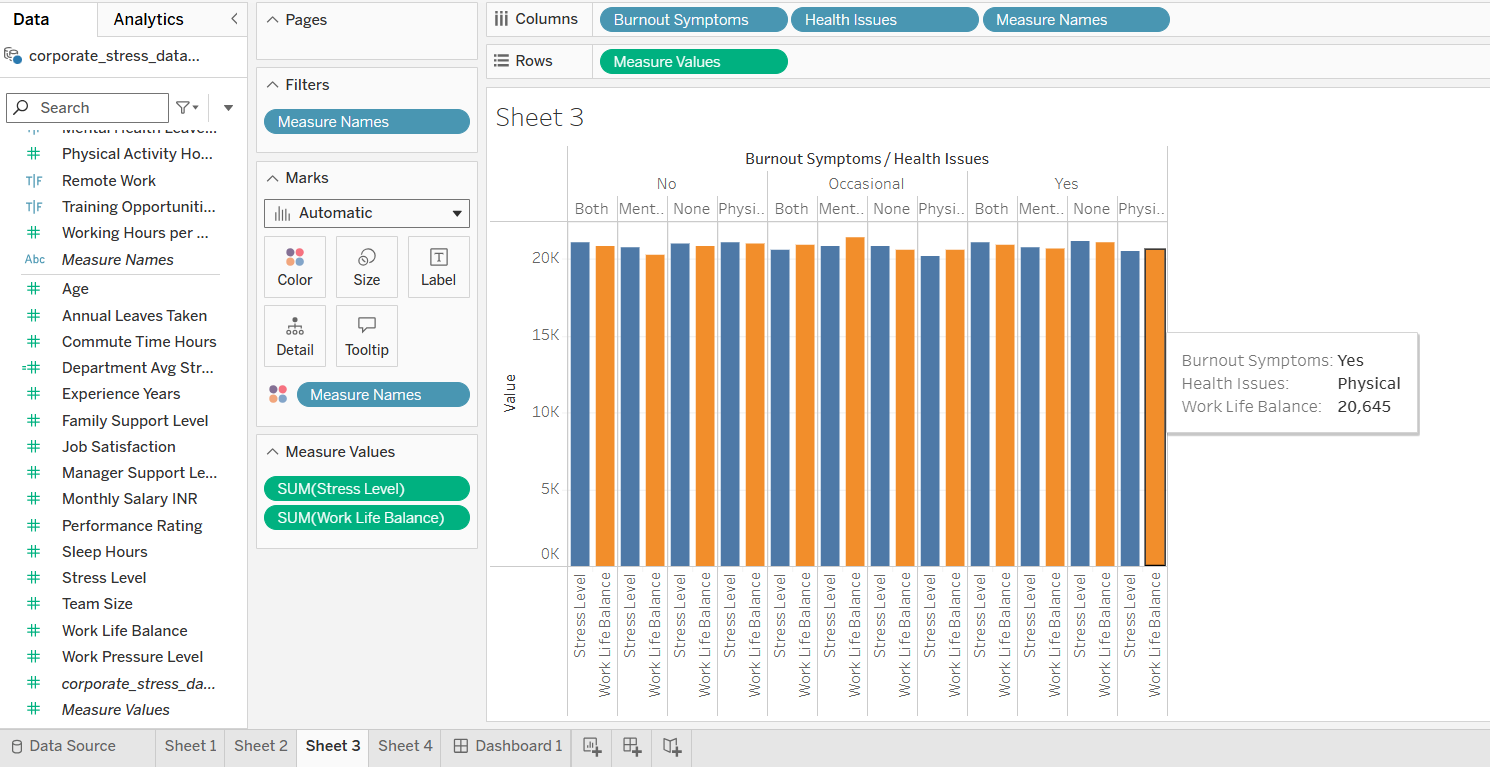




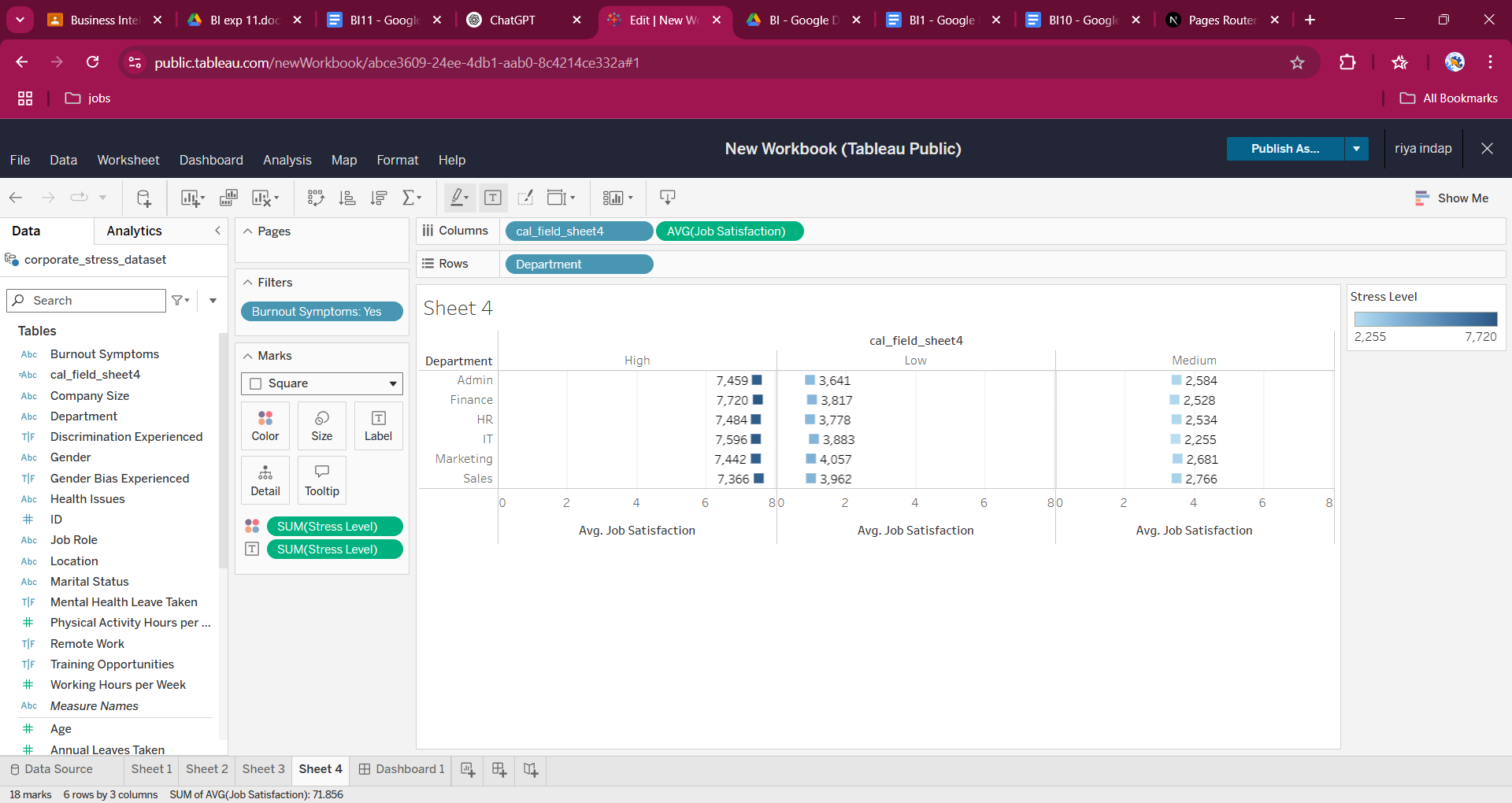
The visualization shows the average stress levels across varying working hours per week, with bar colors indicating total sleep hours. It reveals that as working hours increase, average stress levels remain consistently moderate, but there is a slight fluctuation. Notably, lower sleep hours (darker bars) are associated with slightly higher stress levels, suggesting that inadequate rest could be a contributing factor. However, the trend is not strongly linear, indicating that other factors might also influence stress beyond just working hours and sleep.



The pie chart shows that the Sales department has the highest total stress level, indicating a high-pressure environment despite managerial support. HR also stands out with the highest average stress level, suggesting intense stress among fewer employees. While departments like Marketing and IT show slightly lower stress, overall stress appears consistently high across all departments, highlighting the need for targeted stress management strategies.

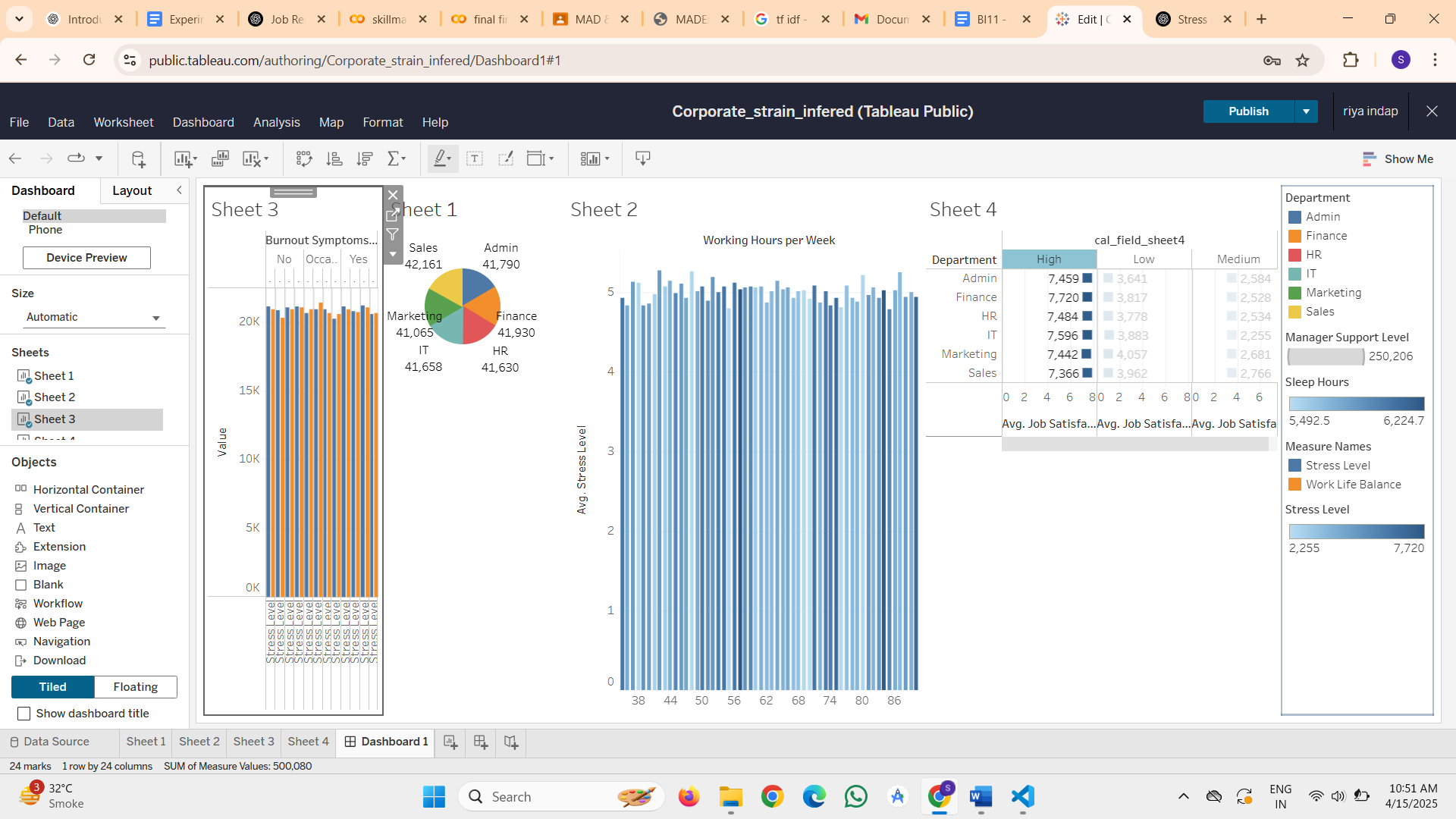


This bar chart compares Stress Level and Work-Life Balance across various combinations of Burnout Symptoms and Health Issues. It shows that stress levels remain consistently high regardless of the type or presence of burnout or health issues. However, those experiencing both burnout and physical health issues tend to have lower work-life balance, highlighting a clear negative impact. Even when burnout symptoms are "occasional" or "none," stress levels still stay elevated, suggesting that stress is a widespread issue and not solely tied to visible burnout or health problems.

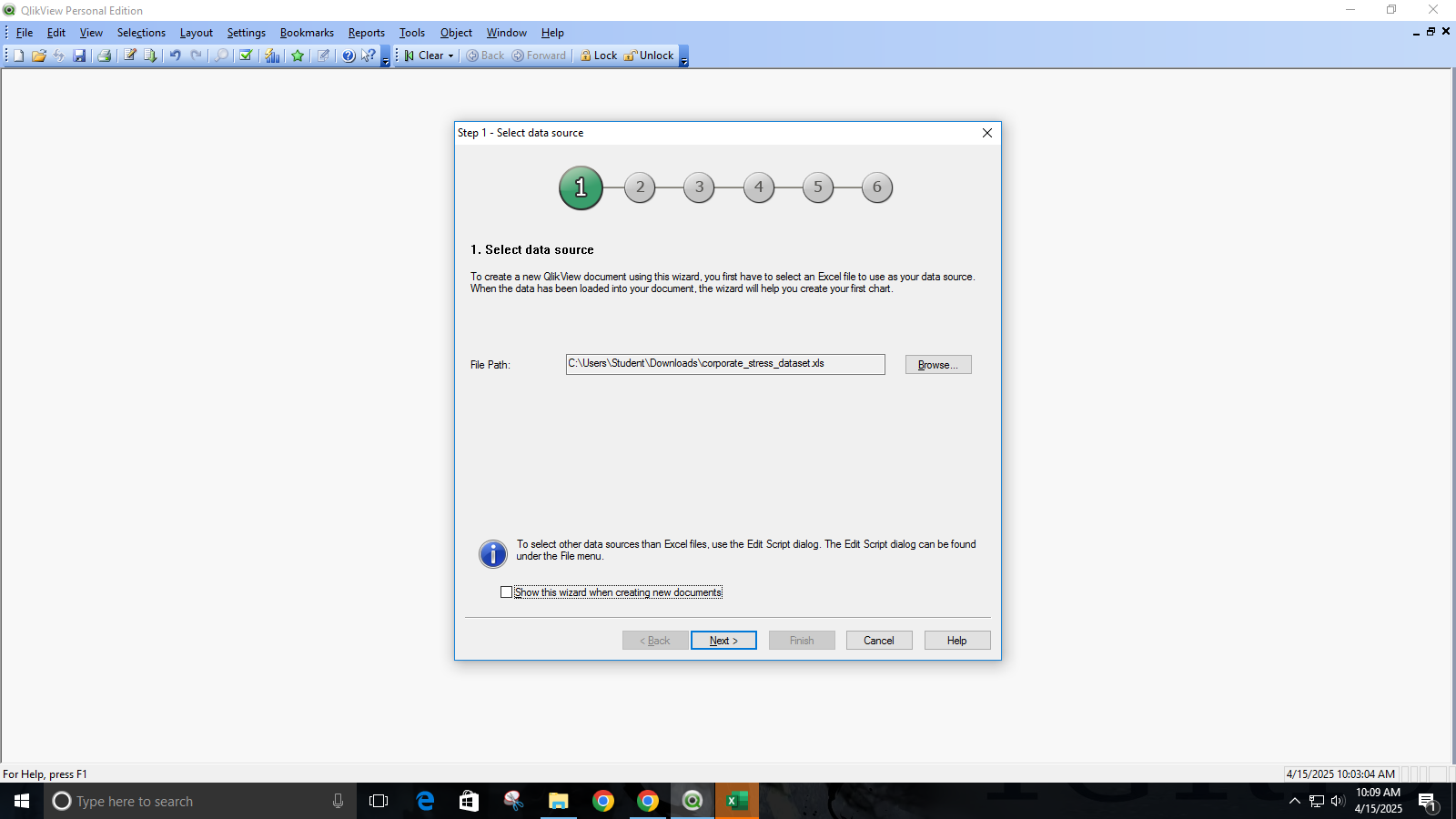


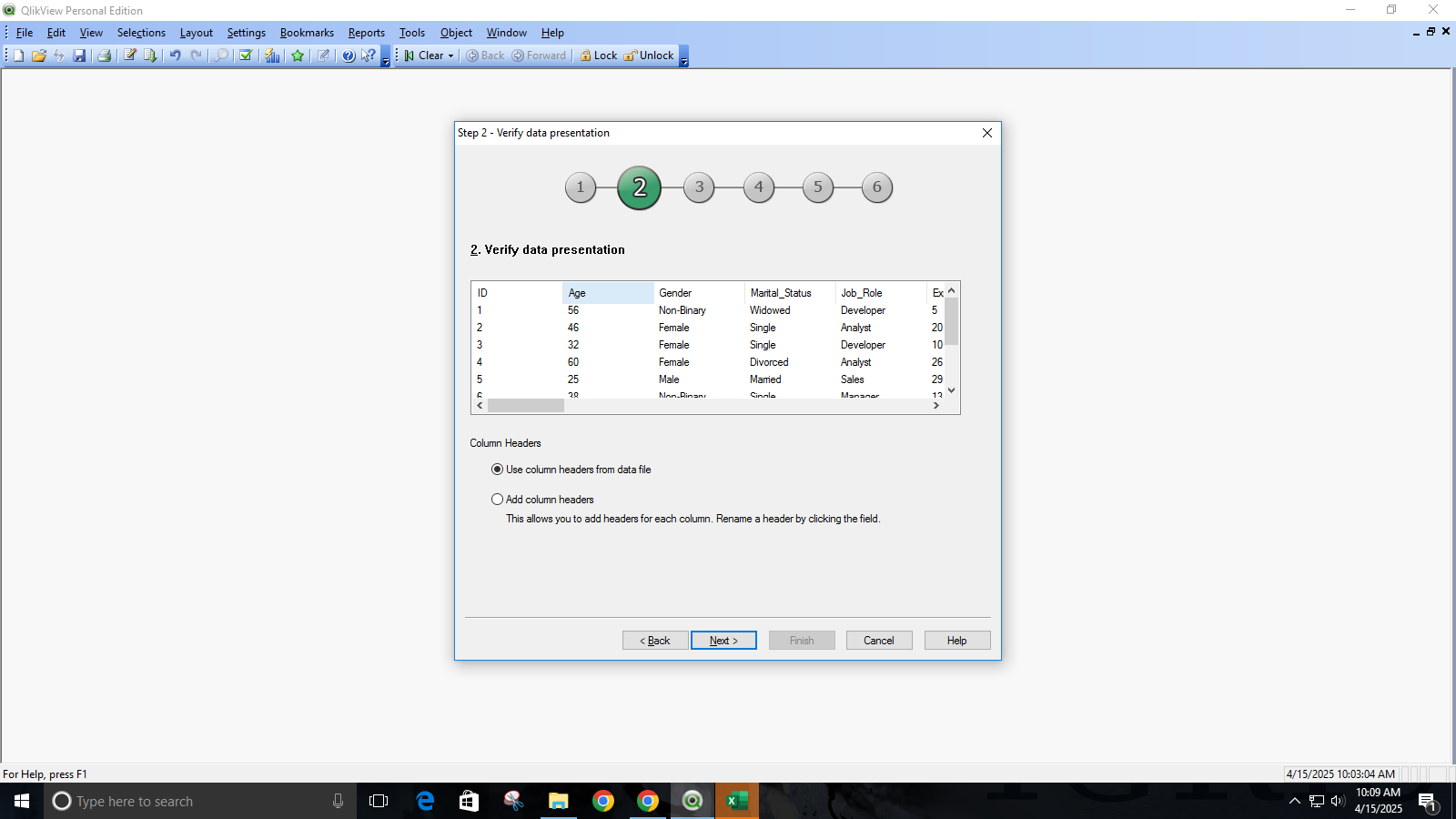
The heatmap shows that departments with high stress levels, like HR and IT, also report high job satisfaction among employees with burnout, suggesting overcommitment. Medium stress departments have the lowest satisfaction, indicating possible lack of support. Even low-stress areas show only moderate satisfaction, highlighting the need for balanced, department-specific interventions.

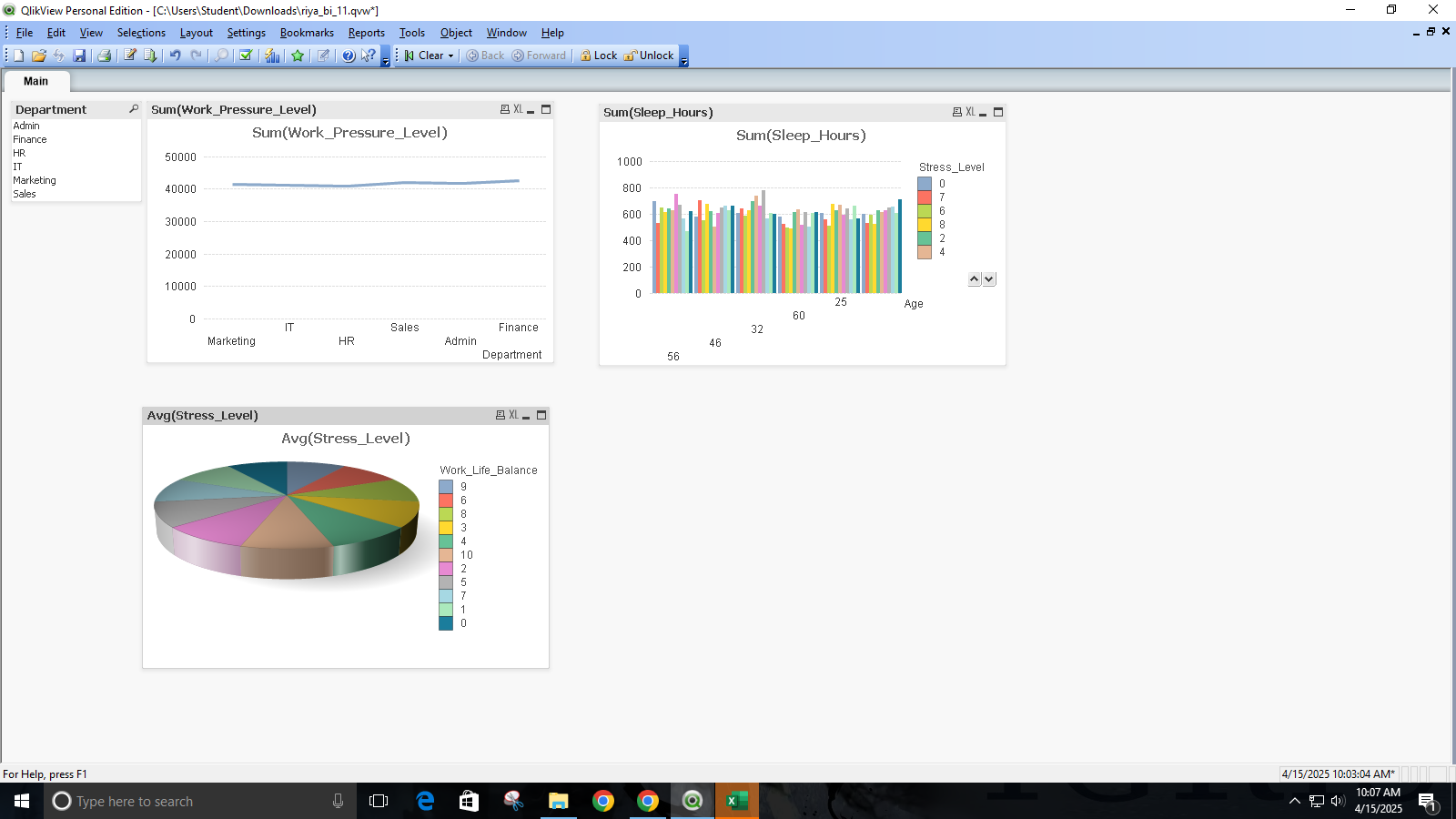
Publishing the story:



Qlikview implementation:







The line chart displaying the Sum of Work Pressure Level across departments shows a relatively consistent yet high level of pressure in all departments, with slight variations, suggesting that work pressure is a common concern across roles.

The bar chart mapping Sum of Sleep Hours against Age and Stress Level highlights that sleep duration remains fairly uniform across age groups, but when segmented by stress levels, a subtle inverse relationship is observed—individuals with higher stress levels tend to have slightly lower sleep hours.

Lastly, the 3D pie chart depicting the Average Stress Level categorized by Work-Life Balance reveals a crucial pattern: employees reporting better work-life balance generally have lower average stress levels. This emphasizes the importance of maintaining a healthy work-life balance to mitigate corporate stress.

Overall, the visualizations collectively underline that stress in modern workplaces is intricately linked to work pressure, insufficient rest, and imbalanced personal-professional boundaries.

**Conclusion:**

**BI Decision:**

1. Prioritize high-risk employees (e.g., burnout clusters) with mental health support and flexible work policies.
2. Train managers to improve support, focusing on high-stress departments (HR, Sales).
3. Reform workloads, limit overtime, and promote remote work for better work-life balance.
4. Launch sleep/wellness programs and use BI dashboards (Tableau/Qlikview) for real-time stress monitoring.

Outcome:

The BI-driven strategies will reduce burnout by targeting high-risk employees with mental health support and flexible policies. Managerial training in high-stress departments (HR/Sales) will improve support satisfaction. Workload audits and remote-work adoption aim to balance hours and boost work-life balance. Department-specific fixes (e.g., role rotation in Sales, agile practices in IT) will lower stress. Sleep campaigns and wellness programs will address stress-sleep links, while BI dashboards (Tableau/Qlikview) enable real-time stress tracking for proactive adjustments. Collectively, this fosters a healthier, more productive workforce with reduced attrition and sustainable organizational resilience.