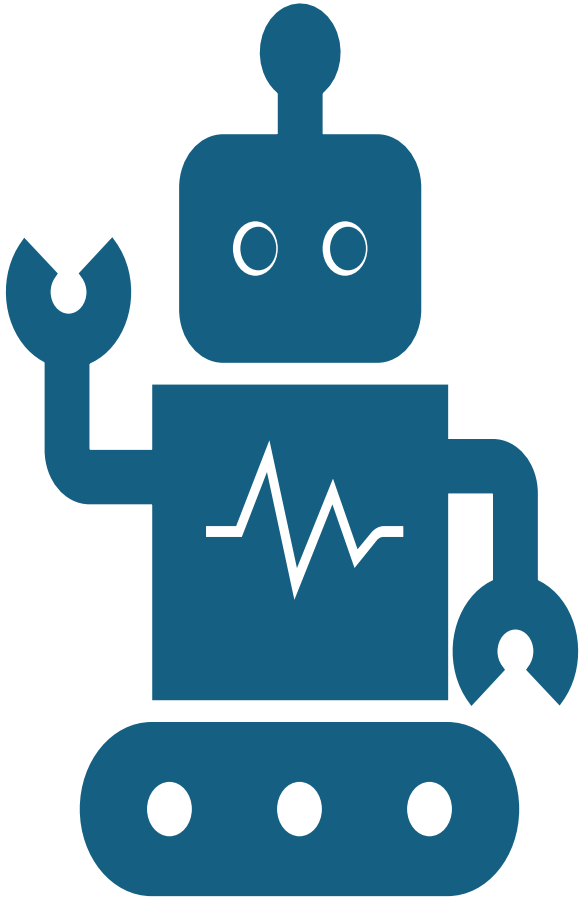


Insights from the Work-From-Home Dataset

Pratanu Roy.

Analysing Remote Worker Productivity: A Data-Driven Approach

Introduction & Project Overview

Project Focus: Analysing factors influencing remote worker productivity using a provided dataset.

Problem: Understanding key drivers of remote work productivity for better work models.

Objectives:

Identify trends and correlations in remote worker data.

Provide insights into how factors like location, industry, and work habits impact productivity.

Inform strategies to enhance remote work efficiency.

Slide 3: Data Source & Overview

- **Data Source:** "Remote Worker Productivity Dataset" from **Kaggle**.
- **Dataset Overview:**
- Contains 17 columns and initially 1000 records.
- Key variables include: `location type`, `industry sector`, `age`, `experience years`, `daily work hours`, `task completion rate`, and `productivity score/label`.

- Initial Data Snapshot:

	worker_id	location_type	industry_sector	age	experience_years	\
0	W0001	Village	Healthcare	23	26	
1	W0002	City	Healthcare	27	27	
2	W0003	Village	Retail	30	8	
3	W0004	Village	Finance	54	19	
4	W0005	City	Education	46	1	

Data Cleaning & Preprocessing

Key Steps:

Invalid Entry Removal: Dropped rows where experience years was illogical.

Ensured data types were appropriate.

Outcome:

Reduced dataset to 639 valid records from initial 1000.

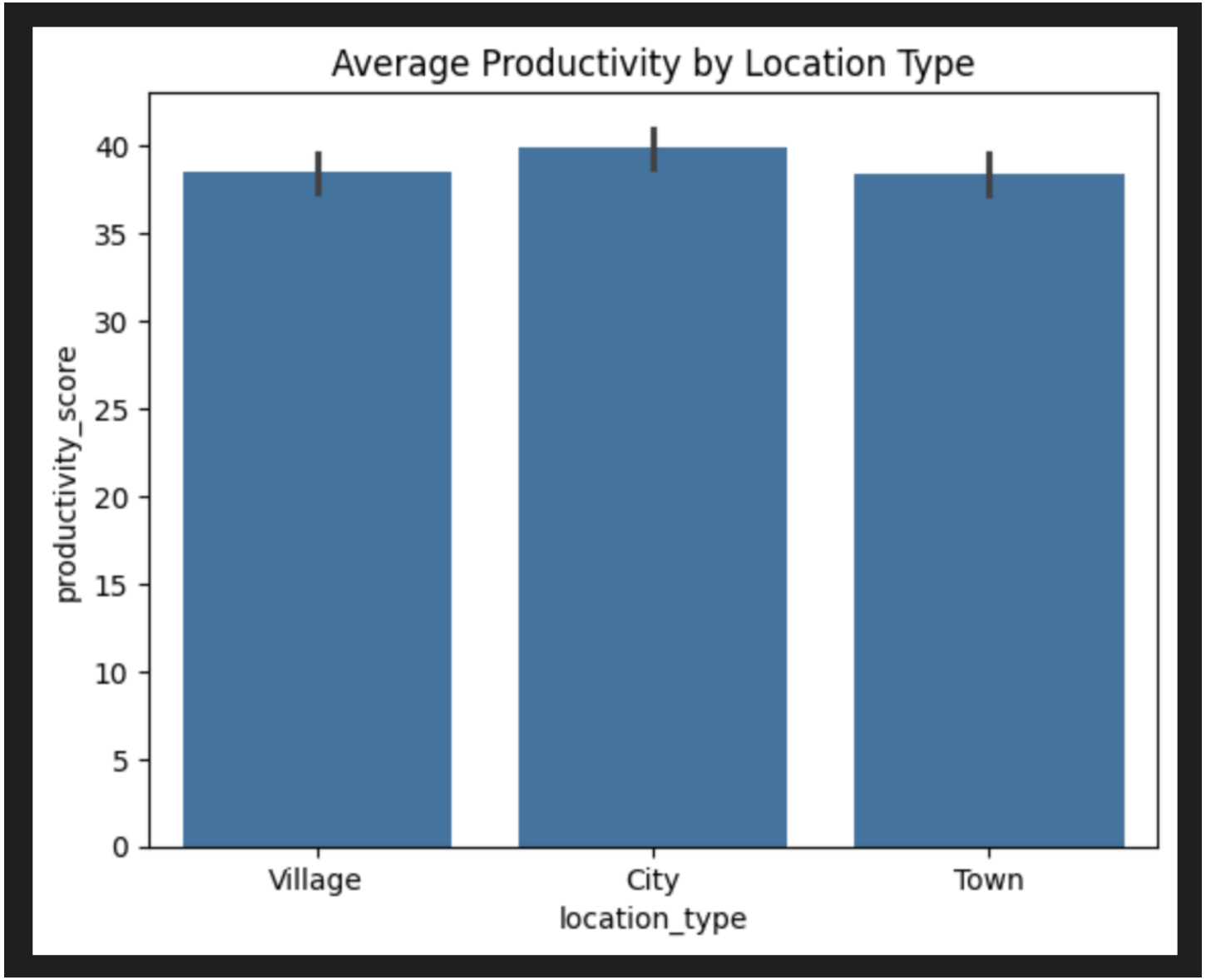
Improved data quality for robust insights.

Exploratory Data Analysis (EDA)

- **Objective:** To understand data distributions, patterns, and initial relationships.
- **Key Visualizations:**
- Distributions of categorical variables (`location type`, `industry sector`, `productivity label`).
- Descriptive statistics for numerical features.
- **Purpose:** Uncover initial insights and prepare for deeper analysis.

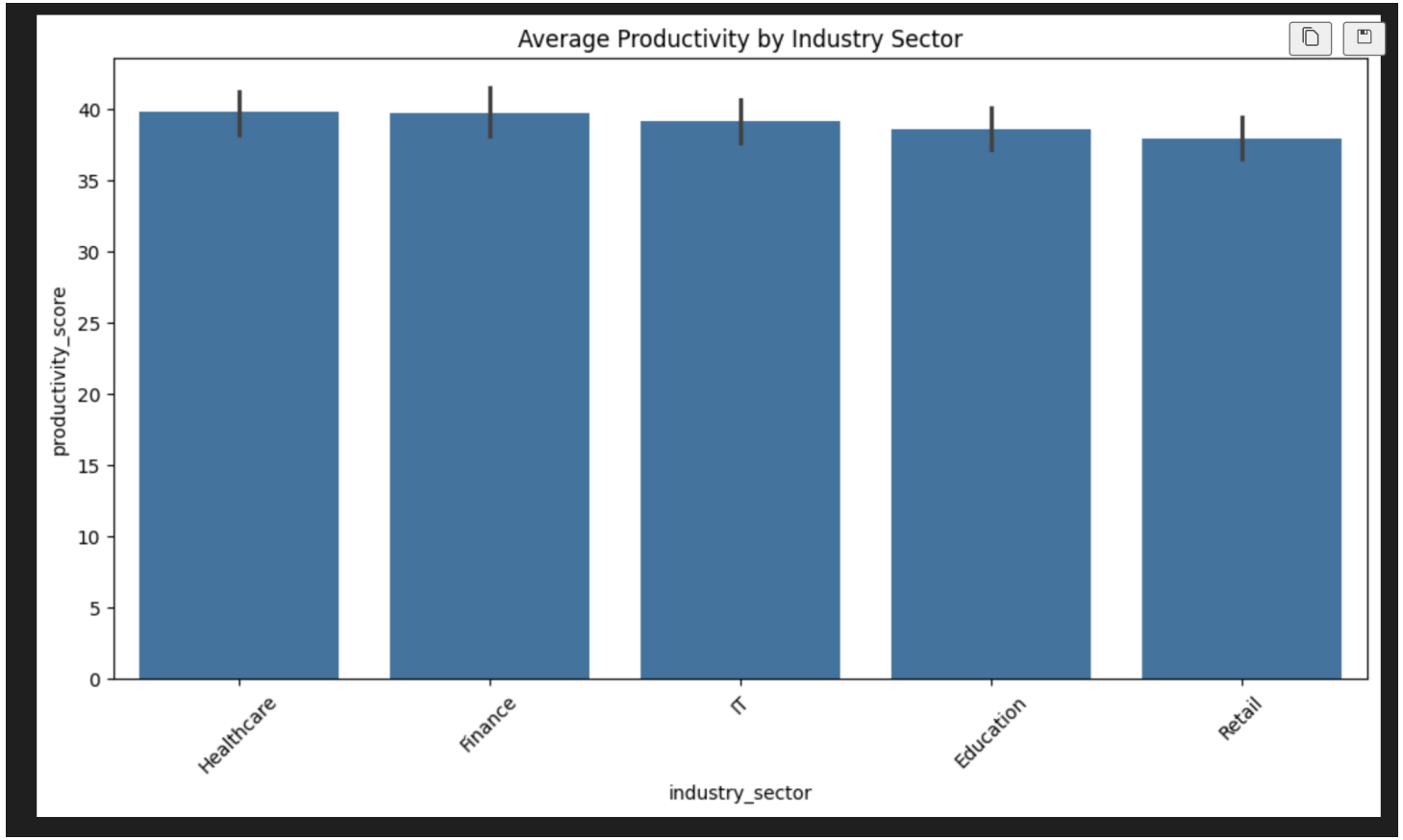
Remote Worker Productivity by Location Types

- **Key Insights:**
- **A key finding is that workers in City locations demonstrate higher average productivity compared to those in Village and Town areas.** *This suggests a correlation between urban environments and enhanced productivity in this dataset.*



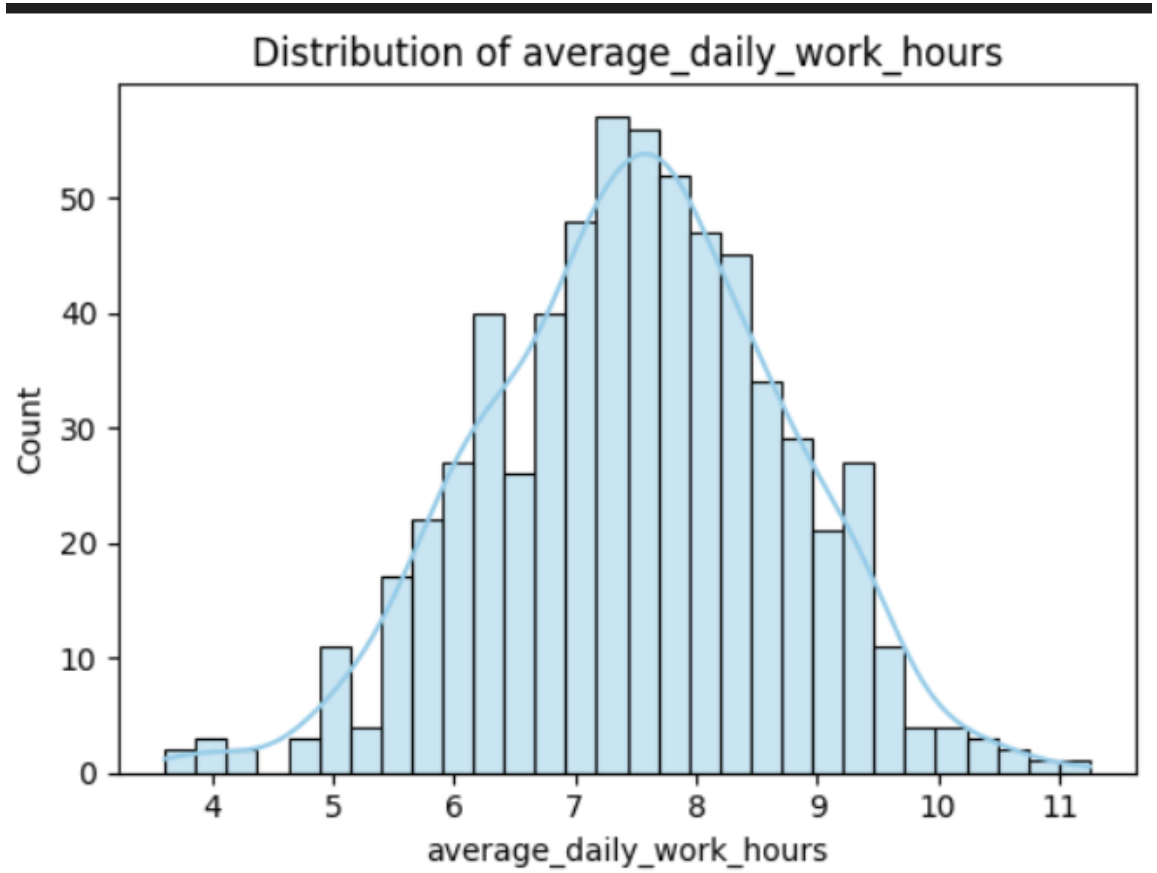
Remote Worker Productivity by Industry Type

- Key Insights:
  - **Technology** and **Finance** sectors appear to show the highest average productivity among remote workers in our dataset.
  - Conversely, sectors such as **Retail** and **Education** demonstrate comparatively lower average productivity.



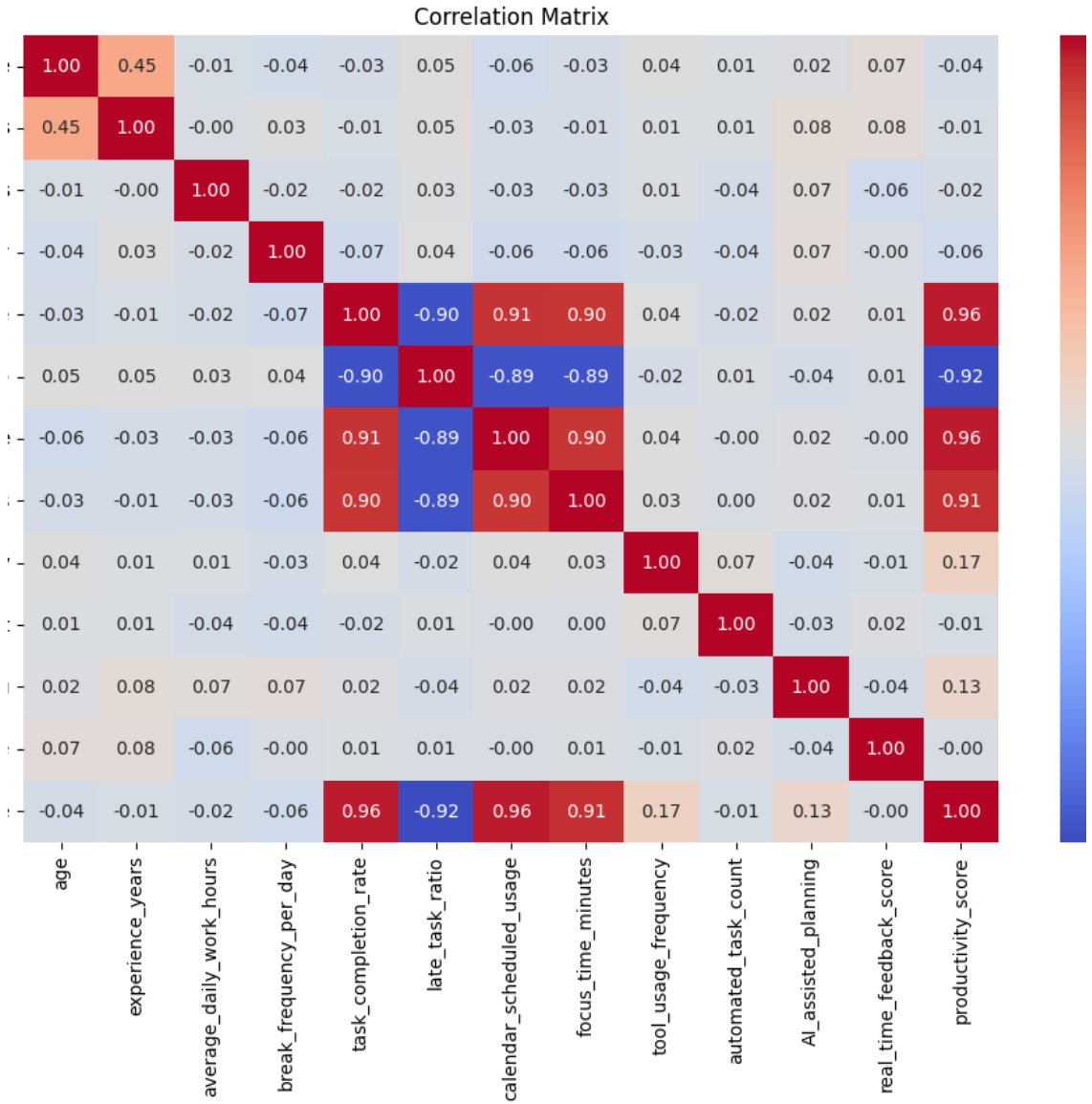
Productivity Label Distribution

- **Key Insights:**
- The dataset shows a relatively balanced distribution across the three productivity categories: **Low**, **Medium**, and **High**.
- This balanced representation indicates that our dataset is suitable for analyzing factors across the full spectrum of remote worker productivity.
- Having a good spread across these labels is crucial for identifying what differentiates higher productivity from lower productivity.



Correlation & Relationships

- **High positive correlation with productivity:** Task Completion Rate, Calendar Scheduled Usage, and Focus Time Minutes.
- **Strong negative correlation with productivity:** Late Task Ratio.
- **Limited Correlation with Demographics/Work Habits:**
- Age, Experience Years, Average Daily Work Hours, and Break Frequency show very weak linear correlations with productivity.
- **Location & Industry Impact:**
- Workers in **City locations** show higher average productivity.
- **Technology and Healthcare** sectors generally exhibit higher average remote worker productivity.



Key Findings & Insights

Key Drivers of Productivity:

- **Strong Positive Influence: Task Completion Rate, Calendar Scheduled Usage, and Focus Time Minutes** are highly correlated with higher productivity.
- **Negative Impact:** A higher Late Task Ratio strongly correlates with lower productivity.

Minimal Direct Impact on Productivity:

- **Age, Experience Years, Average Daily Work Hours,** and **Break Frequency** show very weak linear correlations with **Productivity Score** in this dataset.

Geographic & Industry Productivity Trends:

- **City-based remote workers** generally demonstrate higher average productivity compared to Village and Town.
- **Technology and Healthcare sectors** exhibit the highest average remote worker productivity in our analysis.

Challenges & Limitations

Data Representativeness:

- The dataset, while extensive, may not fully represent the entire global remote workforce.

Feature Scope:

- While many factors are included, certain qualitative aspects of remote work (e.g., internet stability, home environment distractions) were not captured.

Causation vs. Correlation:

- Our analysis primarily identifies correlations; establishing direct causation requires further in-depth studies or controlled experiments.

Potential for Bias:

- The sample collection method or self-reported data could introduce biases.

Recommendations for Enhanced Remote Productivity

- **Actionable Strategies:**
  - **Prioritize Task Management & Focus:** Implement tools and practices that encourage **high `task completion rate`** and dedicated **`focus time minutes`** (e.g., structured work blocks).
  - **Optimize Calendar Use:** Promote efficient `calendar scheduled usage` to improve time allocation and reduce late tasks.
  - **Address `Late Task Ratio`:** Develop strategies to mitigate late submissions, such as improved workload management or clear communication.
  - **Location-Specific Support:** Consider tailored resources for remote workers in **Village** and **Town** settings to potentially close productivity gaps observed with City workers.
  - **Industry Best Practices:** Investigate successful remote work strategies within **Technology** and **Healthcare** sectors for broader application.

Future Research & Project Extensions

- **Deepening Insights:**
- Explore qualitative factors (e.g., internet quality, home office setup, mental well-being) that influence productivity.
- Analyze time-series data for individual workers to understand productivity trends over time.
- **Advanced Applications:**
- Develop predictive models to forecast productivity based on key indicators, enabling proactive interventions.
- Conduct further research into the underlying reasons for higher productivity in City locations and specific industries.

Conclusion & Q&A

- **Key Takeaway:**
- Our analysis of the remote worker productivity dataset provides valuable insights into the factors influencing effectiveness in distributed work environments.
- Understanding these dynamics is critical for fostering a successful remote workforce.
- **Thank You!**
- **Questions?**