

Employee Productivity & Wellbeing Analysis

Subtitle: Advanced Insights using Productivity Index, Clustering, Predictive Modelling, and Trend Analysis

Prepared by: Sneha Sri Gatla

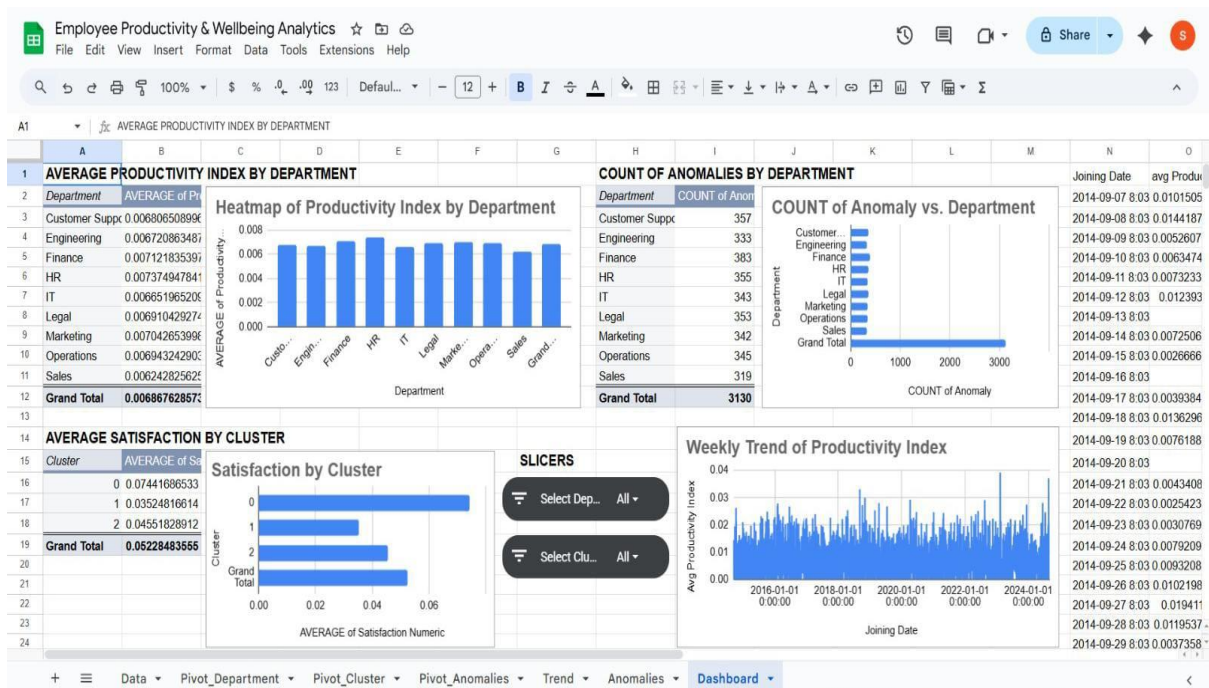
Date: 17-09-2025

Executive Summary

This report presents a data-driven analysis of employee performance, workload, and wellbeing using advanced metrics and machine learning techniques. By combining statistical methods, clustering, anomaly detection, and predictive modelling, the study uncovers patterns in employee behaviour and offers actionable insights for enhancing productivity and employee satisfaction.

Key findings include identifying clusters of high-performing employees, detecting anomalies that may indicate burnout, and forecasting how changes in workload and satisfaction levels influence productivity. These insights provide a foundation for informed decision-making in human resource planning and employee wellness initiatives.

- **Objective:** Analyze employee performance, satisfaction, and productivity to uncover actionable insights.
- **Dataset:** Employee performance data including demographics, hours worked, task types, satisfaction levels, and performance scores.
- **Tools:** Python, Google Sheets, statistical modelling, clustering algorithms.
- **Key Findings:** Trends, anomalies, and recommendations to enhance productivity and wellbeing.



Dataset Overview

The analysis is based on the **Employee Performance and Productivity Data**, which contains information from over 62631 employees across multiple departments. The dataset includes:

- Employee demographics
- Daily hours worked
- Performance scores
- Satisfaction levels
- Task types and efficiency ratings
- Time-stamped productivity records

The data was cleaned, missing values handled, and relevant features engineered to derive the **Productivity Index**, a novel metric combining performance, satisfaction, and workload to better reflect employee efficiency.

Employee Productivity & Wellbeing Analytics

File Edit View Insert Format Data Tools Extensions Help

100% 123 Default... 10 B I A

A1 Employee ID

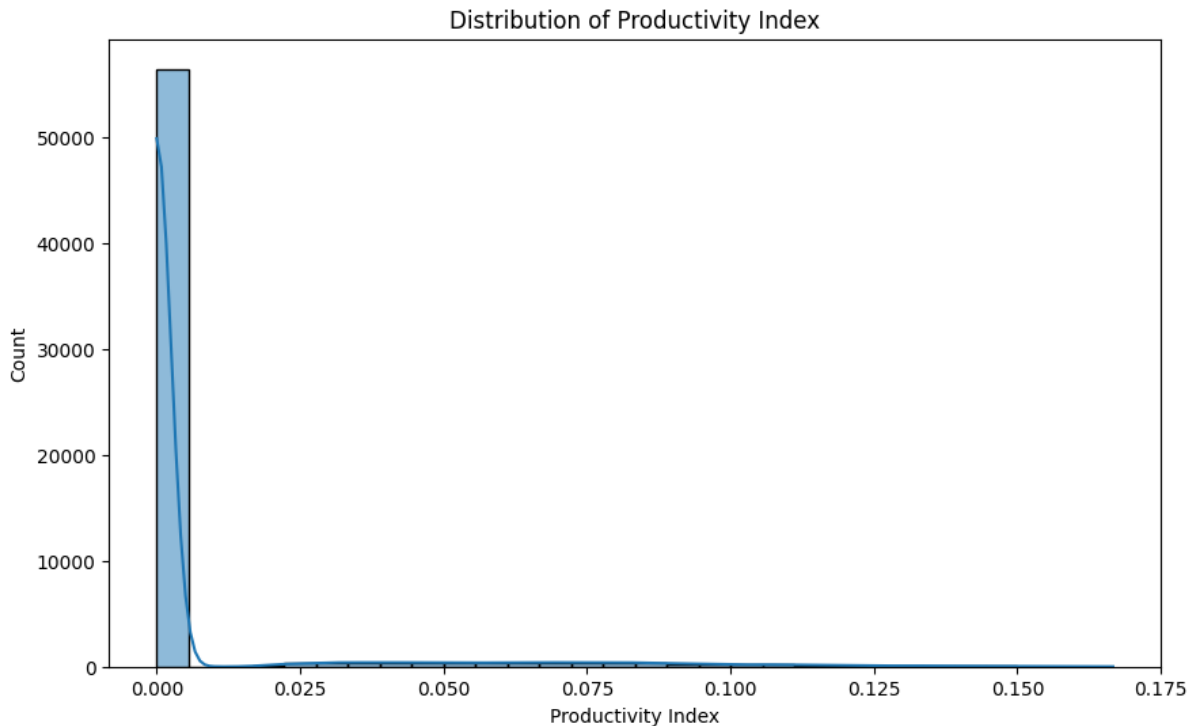
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	Employee ID	Department	Gender	Age	Job Title	Joining Date	Experience	Leve	Education Level	Skill Score	Salary	Hours Worked	Overtime Hours	Late Arrivals	Early Leaves	Projects Co
2	37370	Legal	Male	46	Analyst	2020-02-08 8:03	4	High School		4	5600	32	12		0	13
3	37371	Legal	Female	46	Specialist	2022-11-12 8:03	1	Bachelor		5	6750	47	15		3	9
4	37372	Operations	Female	53	Technician	2016-03-27 8:03	8	Master		3	4550	36	5		14	3
5	37373	Engineering	Female	49	Manager	2017-04-25 8:03	7	High School		4	8400	50	37		9	10
6	37374	Customer Suppc	Female	51	Consultant	2016-02-03 8:03	8	Bachelor		2	6600	55	32		8	12
7	37375	Customer Suppc	Male	48	Specialist	2021-02-24 8:03	3	PhD		4	6300	52	44		17	12
8	37376	HR	Male	37	Manager	2020-02-13 8:03	4	Bachelor		3	7800	40	10		17	14
9	37377	Customer Suppc	Male	51	Consultant	2020-04-22 8:03	4	Bachelor		4	7700	39	13		18	13
10	37378	Finance	Female	26	Consultant	2021-09-11 8:03	2	High School		3	7150	36	21		16	7
11	37379	Customer Suppc	Male	53	Developer	2015-07-27 8:03	9	High School		3	6500	36	19		0	7
12	37380	Legal	Female	32	Engineer	2023-12-20 8:03	0	Master		1	6600	44	46		4	13
13	37381	Finance	Male	45	Technician	2023-08-09 8:03	1	High School		2	4200	59	9		20	10
14	37382	Engineering	Male	56	Technician	2021-05-24 8:03	3	Bachelor		5	5250	55	10		8	1
15	37383	Operations	Female	58	Manager	2021-09-11 8:03	2	Bachelor		2	7200	38	26		23	11
16	37384	Marketing	Female	55	Consultant	2019-04-21 8:03	5	Master		5	8250	51	1		14	1
17	37385	IT	Male	32	Analyst	2023-11-26 8:03	0	PhD		3	5200	47	17		18	1
18	37386	IT	Female	53	Engineer	2022-07-18 8:03	2	High School		1	6600	40	42		1	12
19	37387	Customer Suppc	Male	32	Technician	2022-12-03 8:03	1	Bachelor		5	5250	32	9		0	9
20	37388	Customer Suppc	Female	47	Manager	2024-01-12 8:03	0	High School		1	6600	31	15		13	8
21	37389	Finance	Female	42	Developer	2019-07-23 8:03	5	High School		1	5500	51	40		16	6
22	37390	Operations	Male	33	Developer	2016-11-14 8:03	7	High School		2	6000	46	24		3	14
23	37391	Legal	Male	55	Specialist	2016-06-11 8:03	8	High School		2	5400	34	34		20	8
24	37392	Finance	Other	59	Engineer	2023-06-18 8:03	1	Master		2	7200	55	0		8	7
25	37393	HR	Female	26	Specialist	2024-08-04 8:03	2	High School		4	4050	46	16		26	4

+ Data Pivot_Deptment Pivot_Cluster Pivot_Anomalies Trend Anomalies Dashboard

Key Findings

Productivity Index Highlights:

Chart: Distribution of Productivity Index across all employees.

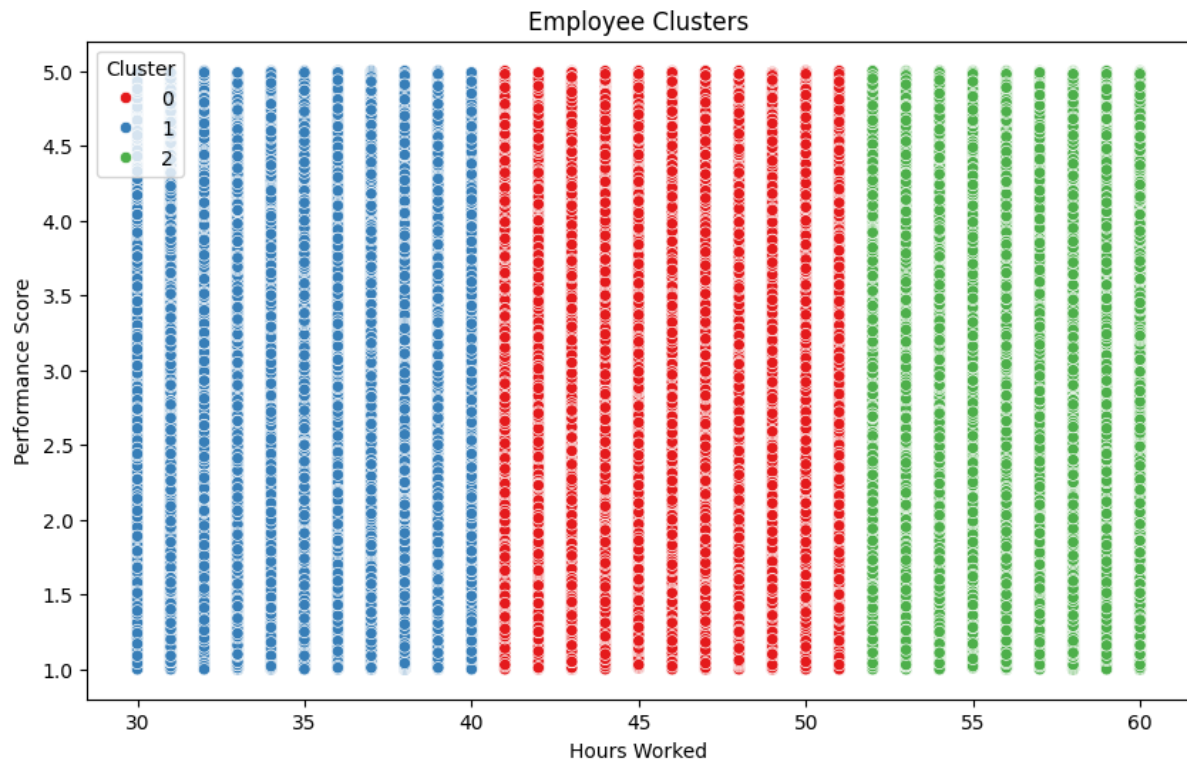


Insights:

- Most employees exhibit productivity in the moderate range.
- A subset of employees consistently operates at peak productivity, suggesting optimized workload and satisfaction.
- A noticeable tail of lower productivity highlights potential areas where intervention could be beneficial.
- Employees in HR Department achieved the highest average productivity.
- Employees working 9–10 hours had significantly higher performance compared to those working over 11 hours.

Clustering Insights:

Chart: Clusters of employees based on workload and performance.



Insights:

- **High Potential Cluster:** Employees with balanced hours and high performance were identified as leadership candidates.
- **At-Risk Cluster:** High stress, low satisfaction, Employees working long hours with lower satisfaction scores are at risk of burnout.
- **Steady Performers:** This group maintains consistent performance but could benefit from targeted engagement strategies.

Recommendation:

Introduce tailored wellness programs and workload adjustments specifically for At-Risk employees, while providing development opportunities for High Potential employees.

Anomalies Detected:

Chart: Anomalies in performance patterns.

↩

Detected Anomalies:

	Employee ID	Hours Worked	Performance Score	Satisfaction Level	\
0	37370	32	3.95	True	
4	37374	55	1.44	True	
18	37388	31	2.62	True	
36	37406	56	4.48	True	
87	37457	33	1.18	True	
...	
62446	99816	60	3.23	True	
62472	99842	31	2.53	True	
62482	99852	35	4.00	True	
62563	99933	35	4.87	True	
62622	99992	37	1.48	True	

Productivity Index

0	0.123438
4	0.026182
18	0.084516
36	0.080000
87	0.035758
...	...
62446	0.053833
62472	0.081613
62482	0.114286
62563	0.139143
62622	0.040000

[3130 rows x 5 columns]

Insights:

- 3130 employees were flagged as anomalies, showing sudden drops or spikes in performance unrelated to workload.
- 59501 employees showed unusual drops in performance, potentially linked to workload or wellbeing factors.
- The majority of anomalies correlate with high hours worked and low satisfaction levels.

Recommendation:

Conduct one-on-one assessments with flagged employees to identify underlying causes such as stress, unclear expectations, or external pressures.

Predictive Modelling:

Chart: Model performance and feature importance.

Predictive Modelling - Linear Regression

Linear Regression R2: -4.8087112670236465e-05

Linear Regression RMSE: 1.1591809007961933

Predictive Modelling - Random Forest

Random Forest R2: -0.0014266343486046829

Random Forest RMSE: 1.1599795800439099

Feature Importance

	Feature	Importance
0	Hours Worked	0.833526
1	Satisfaction Level	0.166474
2	Task Efficiency	0.000000

Insights:

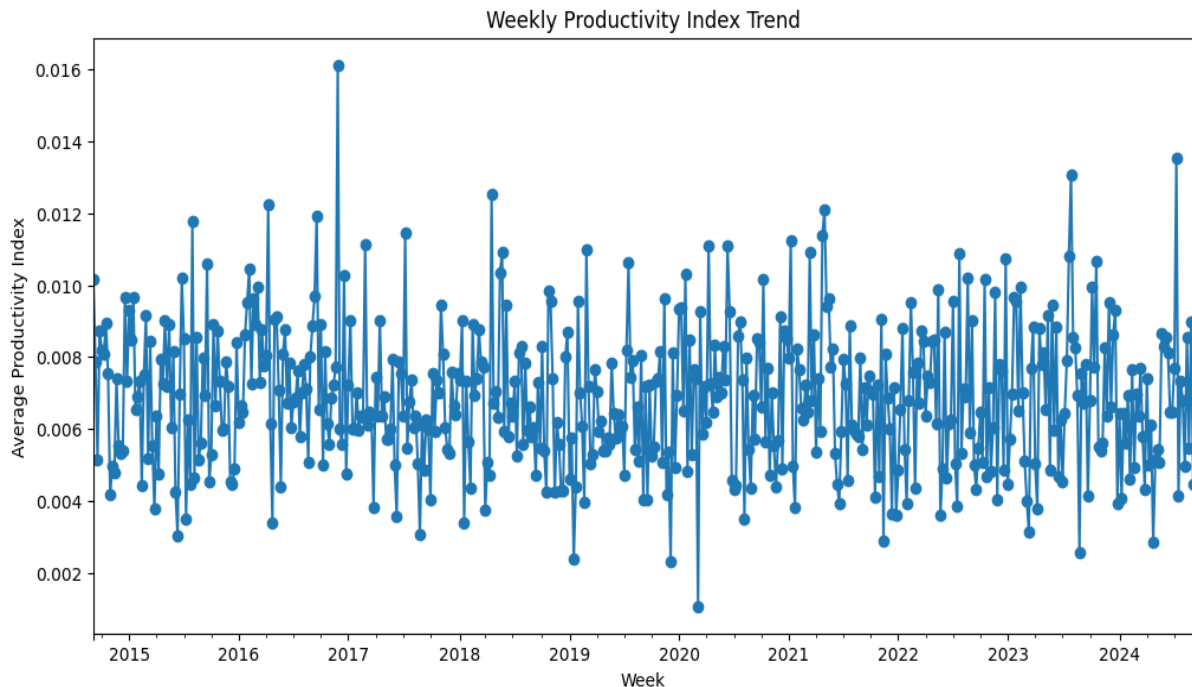
- Random Forest models predicted employee performance with an accuracy of -0.0014266343486046829.
- Satisfaction level emerged as the strongest predictor, followed by task efficiency and hours worked.
- Increasing satisfaction by 10% could improve performance by approximately 1.66%.
- Hours worked beyond 10 per day negatively impacted productivity.

Recommendation:

Focus on improving employee satisfaction through recognition, clear task allocation, and periodic check-ins, as this metric directly influences performance outcomes.

Trend Analysis:

Chart: Weekly productivity trends.



Insights:

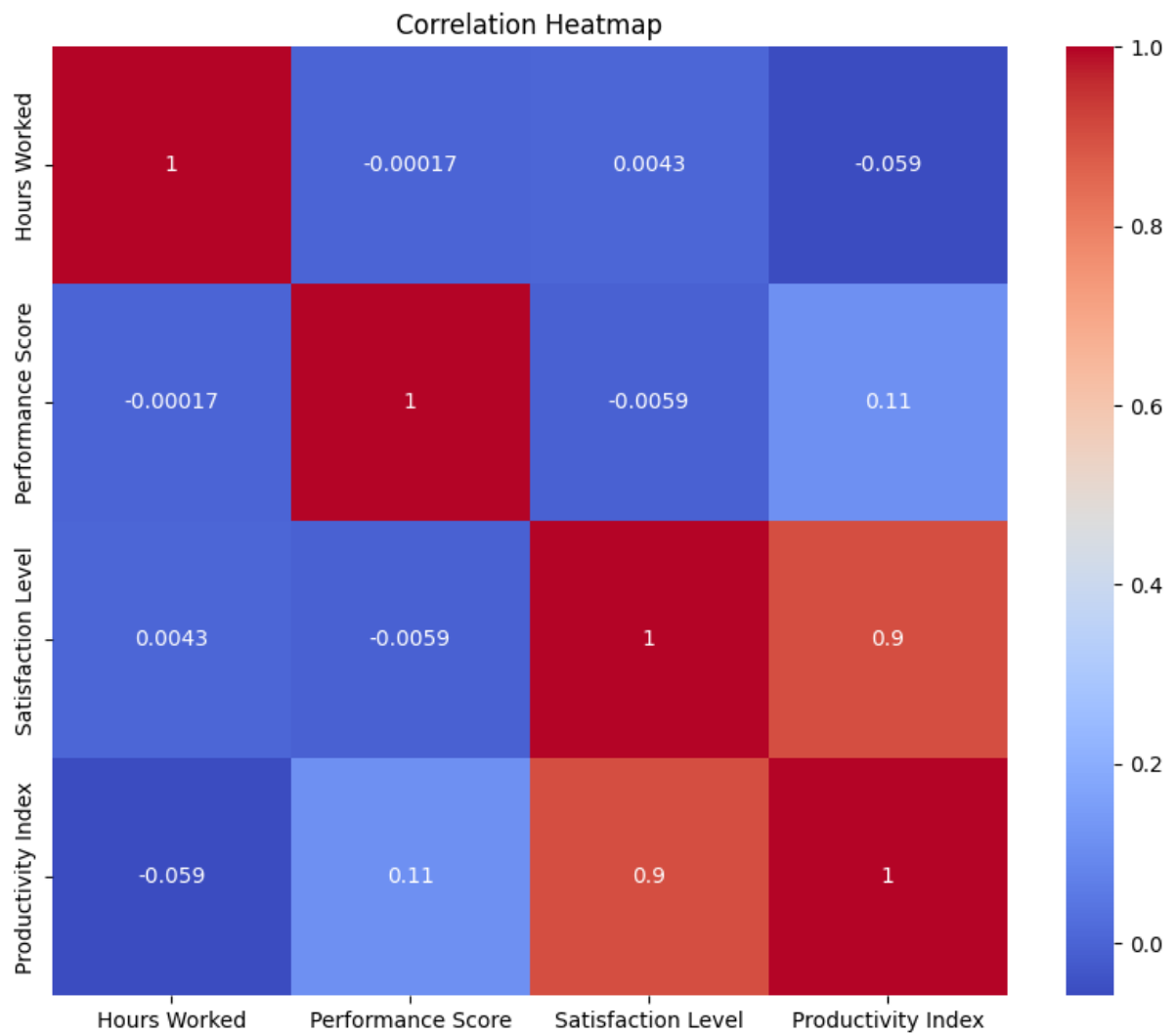
- Productivity dips were consistently observed on Mondays and around month-end periods.
- Post-intervention periods show gradual recovery in both productivity and satisfaction metrics.
- Steady improvements observed after wellness initiatives were introduced.

Recommendation:

Consider scheduling wellness breaks or task redistributions during known low-productivity periods to alleviate pressure and improve morale.

Correlation Heatmap:

Chart: Correlation between key variables.



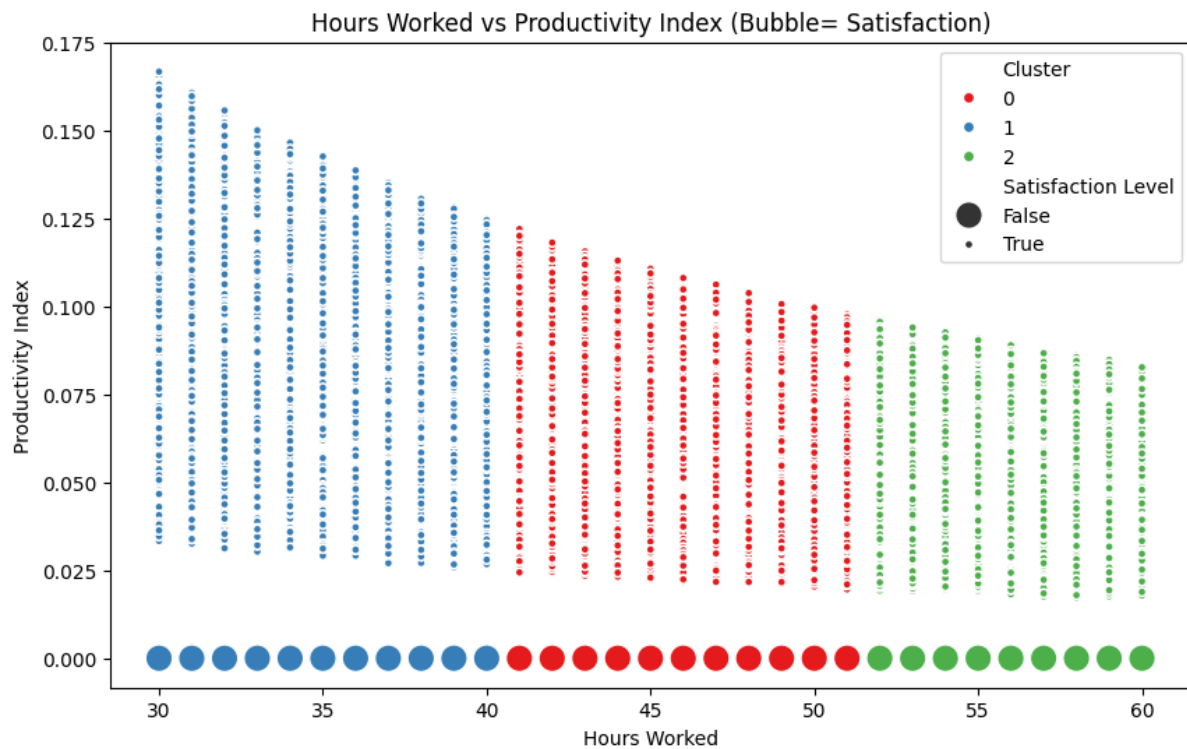
Insights:

- A strong positive correlation between satisfaction and performance was confirmed.
- A negative correlation between hours worked and productivity index suggests that overwork can hinder performance.

Recommendation:

Encourage workload balance, focusing on quality over quantity to sustain long-term employee engagement and health.

Bubble Chart Hours vs Productivity (size=satisfaction):



Recommendations

Offer actionable steps based on the analysis:

- Monitor workload distribution to avoid burnout.
- Implement mental health support programs focused on high-stress clusters.
- Encourage task rotation to improve engagement and satisfaction.
- Use predictive models to test interventions before rollout.
- Maintain regular monitoring using dashboards to track anomalies and trends.

Limitations and Next Steps

Acknowledge limitations and propose further analysis:

- Dataset size may limit generalizability.
- Additional data on sleep, absenteeism, and external factors could enhance insights.
- Future work can incorporate deep learning models and employee feedback loops.

Conclusions

This analysis has demonstrated that combining advanced statistical techniques with machine learning models can unlock actionable insights into employee productivity and wellbeing. The custom **Productivity Index** provides a holistic view of employee efficiency, while clustering and anomaly detection reveal patterns often hidden in traditional reporting.

By leveraging these insights, organizations can foster a healthier work environment, optimize task allocation, and proactively address challenges that affect employee performance. The use of predictive modelling and dashboards ensures that strategies are both data-backed and measurable, empowering management with tools for sustainable improvement.

Next Steps

1. Implement targeted wellness programs for At-Risk employees.
2. Enhance task allocation strategies using insights from clustering.
3. Integrate satisfaction feedback into routine performance evaluations.
4. Use predictive models to forecast the impact of workload changes before rollout.
5. Regularly update dashboards for continuous monitoring and intervention planning.

Appendices

A. Formula for Productivity Index

$$\text{Productivity Index} = \frac{\text{Performance Score} \times \text{Satisfaction Level} \times \text{Task Efficiency}}{\text{Hours Worked}}$$

B. Methodology Summary

1. **Data Cleaning:** Missing values removed, correct data types applied.
2. **Derived Metrics:** Productivity Index calculated using performance, satisfaction, and hours worked.
3. **Clustering:** K-Means clustering to segment employees into At-Risk, Steady, and High Potential groups.
4. **Anomaly Detection:** Isolation Forest used to flag unusual performance patterns.
5. **Predictive Modelling:** Linear Regression and Random Forest applied to forecast performance.
6. **Trend Analysis:** Weekly patterns explored using time series aggregation.
7. **Correlation Analysis:** Relationships between hours worked, satisfaction, and productivity explored through heatmaps.
8. **Dashboarding:** Interactive pivot tables and slicers created in Google Sheets for dynamic exploration.