DSA210

Trading vs Weather Analysis

Author: Atakan Vardar
Date: January 2025

Introduction

This report explores whether personal trading performance correlates with local Istanbul weather conditions (temperature, precipitation, humidity) and time-based factors (weekday, hour of day, monthly patterns). The aim is to see if external factors can help explain or predict daily trading profit/loss (P/L).

Data Sources & Preprocessing

Data Overview

Trading Data

Files: `trading_results_summary.csv`

Columns:

- `Open_Time`, `Close_Time` (DateTime)
- `Profit` (daily net P/L)
- `Result` (Win/Lose/BreakEven)
- Risk and other columns (Stop_Loss, Take_Profit, etc.)

Weather Data

File: `Istanbul,Turkey 2023-08-18 to 2024-12-18.csv`

Columns:

- `Date`
- `temp`, `tempmax`, `tempmin`, `precip`, `humidity`
- Other daily weather indicators (e.g., conditions like “Rainy”, “Cloudy”).

Both sets are merged on a daily `Date` to align trading outcomes with weather data.

Preprocessing Steps

Date Conversion:

- Extracted daily `Date` from `Open_Time` (trading).
- Renamed `datetime` → `Date` in weather and normalized times to `00:00`.

Feature Engineering:

- `rain_indicator` = 1 if `precip` > 0, else 0
- `temp_range` = `tempmax - tempmin`
- `day_of_week` (0=Monday, ..., 6=Sunday)
- `hour_of_day` (0–23, from `Open_Time`)
- `trade_duration_min` = (Close_Time − Open_Time) in minutes

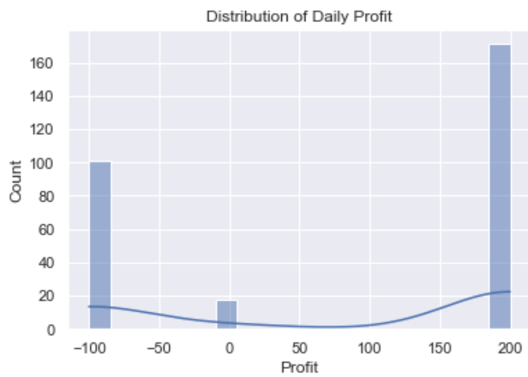
Merged Dataset

After merging on `Date` with an **inner join**, the final dataset has **290 rows** of daily records, each containing:

- **Trading**: Profit, Result, Open_Time, Close_Time, etc.
- **Weather**: Temperature, precipitation, humidity, etc.

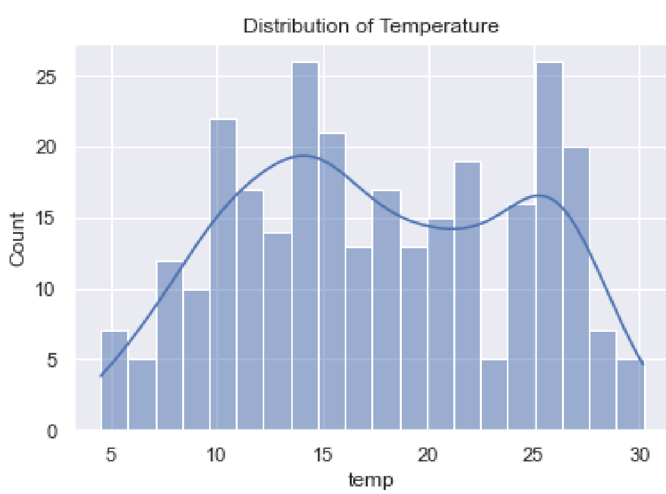
Exploratory Data Analysis (EDA)

Profit Distribution



- Figure: *Distribution of Daily Profit* (Histogram)
- Shows strong peaks at approximately -100 and +200, suggesting frequently used stop-loss and take-profit levels.
- A smaller cluster near 0 indicates break-even trades.Weather Variables

Temperature Distribution



- Figure: *Distribution of Temperature*
- Ranges ~5°C to 30°C, with a typical peak around ~15–16°C

Profit on Rainy vs. Non-Rainy Days

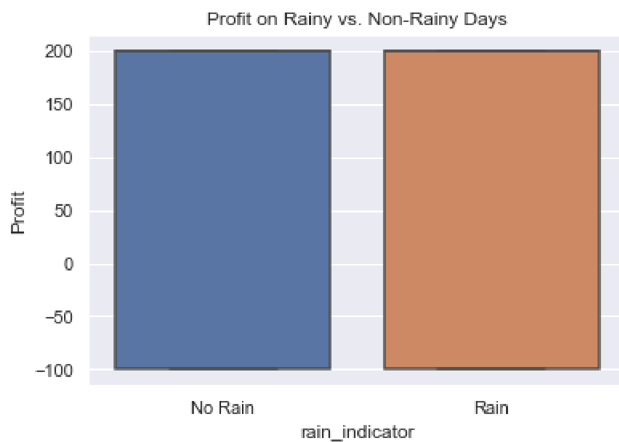


Figure: *Profit on Rainy vs. Non-Rainy Days* (Boxplot)

- Shows no obvious difference in average profit when comparing rainy vs. non-rainy days.

Correlation Matrix

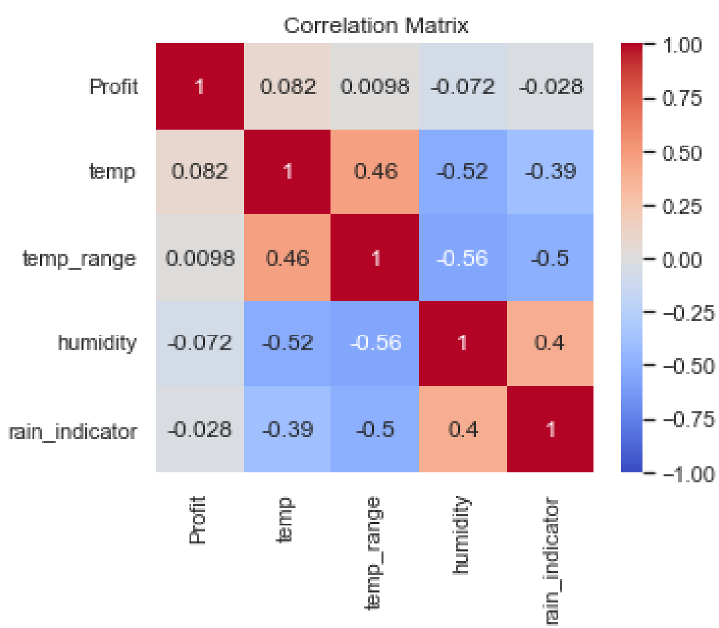
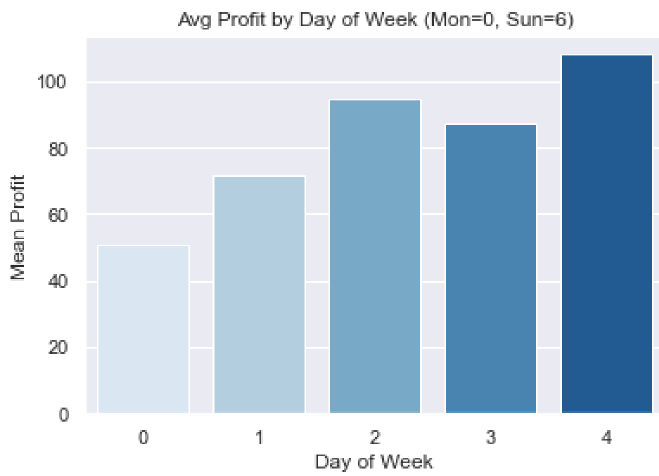


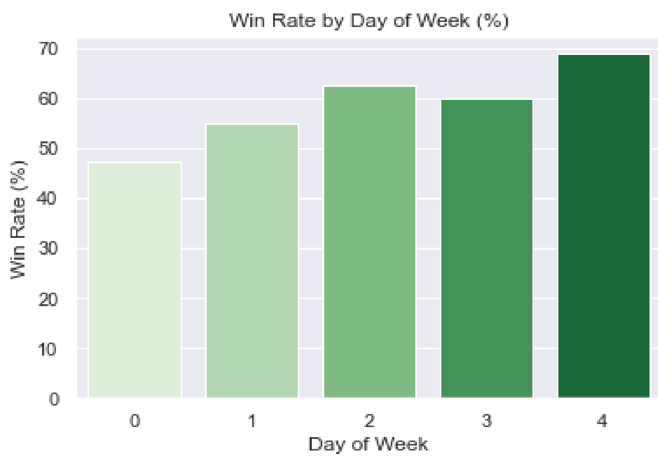
Figure: *Correlation Matrix (Profit, temp, humidity, rain_indicator, etc.)*

- Weak correlations between `Profit` and weather features.

Time-Based Analysis (Daily)

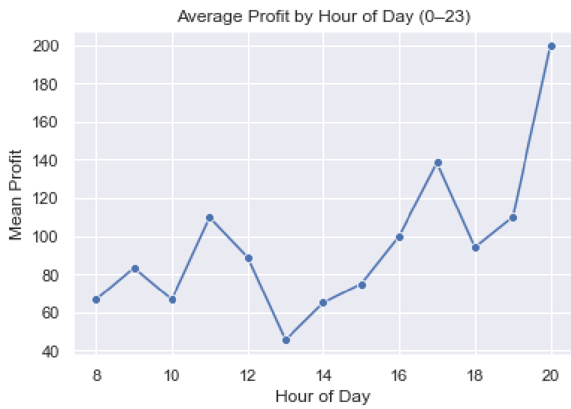
By Day of Week

- 

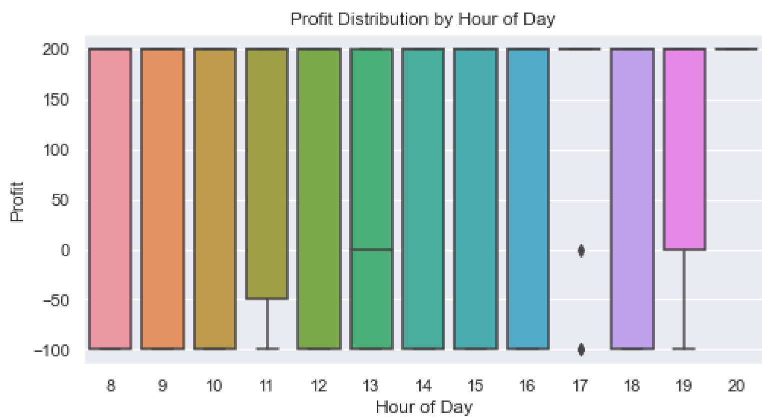
- **Figure**: *Avg Profit by Day of Week*
- 

- **Figure**: *Win Rate by Day of Week*
- Observations: Friday (day=4) has the highest average profit (~100+). Monday (day=0) is lower (~40–50). Friday's win rate is ~70%.

By Hour of Day

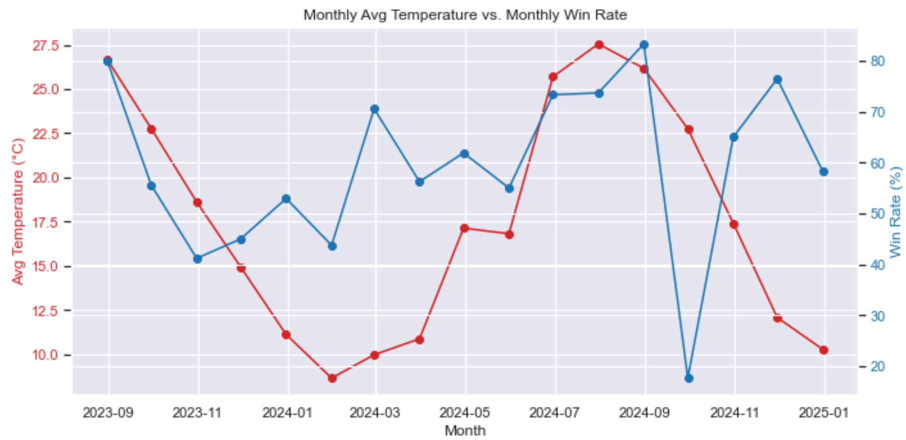


-
- **Figure**: *Average Profit by Hour of Day* (line chart)



-
- **Figure**: *Profit Distribution by Hour of Day* (boxplot)
- Observations: Late evening trades (18–20) show a significant jump in average profit. Midday hours around noon can fluctuate or dip.

Monthly Analysis (Temperature vs. Win Rate)



- **Figure**: *Monthly Avg Temperature vs. Monthly Win Rate*
    - Lines show monthly average temperature (°C) vs. monthly win percentage.
    - No clear strong parallel or inverse pattern is evident.

Statistical Tests

Rain vs. Non-Rain

- H_0: No difference in average Profit on rainy vs. non-rainy days.
- H_1: *There **is** a difference in average Profit on rainy vs. non-rainy days.*

Day of Week

- H_0: Mean Profit is the same across weekdays (Mon–Sun).
- H_1: *Mean Profit **is not** the same across all weekdays; at least one weekday differs significantly.*

Monthly Temp vs. Win Rate

- H_0: No linear or monotonic correlation between monthly temperature and win rate.
- H_1: *There **is** a linear or monotonic correlation between monthly temperature and win rate.*

Rain vs. Non-Rain (t-test)

- *t-statistic* ≈ -0.475, *p* ≈ 0.6355
- $p > 0.05$ → fail to reject H_0. No significant difference found.

ANOVA & Kruskal–Wallis (Day of Week)

- *ANOVA*: F-stat ≈ 1.3971, p ≈ 0.2350
- *Kruskal–Wallis*: H-stat ≈ 5.4153, p ≈ 0.2473
- Both $p > 0.05$ → **fail to reject** H_0. No statistically significant weekday effect.

Monthly Temperature vs. Win Rate (Pearson & Spearman)

- *Pearson Corr* ≈ 0.2641, p ≈ 0.3057
- *Spearman Corr* ≈ 0.3824, p ≈ 0.1299
- Both $p > 0.05$ → No significant correlation.

Simple Regression

A **Linear Regression** model predicted `Profit` from `[temp, humidity, temp_range, rain_indicator, day_of_week, hour_of_day]`.

- **R^2** ~ 0.033 (3.3% variance explained)
- **RMSE** ~ 141.17

Interpretation: Weather/time features alone do **not** strongly predict daily profit.

Key Findings & Interpretations

Rain vs. Non-Rain

- Statistical tests show no significant difference in average daily profit.

Day of Week

- Visual patterns (e.g., Friday looks higher) are **not** statistically confirmed.

Monthly Temp vs. Win Rate

- Pearson/Spearman correlations are > 0.05, so no strong evidence of a relationship.

Regression

- Weather/time features alone yield low predictive power for daily P/L.

Conclusions & Future Work

Weak Weather Impact

- Temperature, rain, humidity do not appear to systematically affect daily profit in this dataset.

Time Factors

- While certain days/hours (like Friday or 8 PM) look better on average, significance tests do not support a robust effect.

Future Directions

- Longer Data Window: Increase sample size across different seasons/years.
- Market Indicators: Add volatility indices, news events, or macro data to see if they overshadow weather/time influences.
- Intraday Analysis: Investigate more granular intervals (e.g., 15–30 minute windows) for micro-trends.
- Advanced Models: Consider time-series forecasting (ARIMA, LSTM) or ensemble methods with more features.