

# Data Mining

## Homework 1

Ashkan Ansarifard

1970082

## Problem 1

1. I choose the probability space such that each element in  $\Omega$  corresponds to a unique ordering of the cards in the deck.

Mathematically,  $\Omega = \{\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_{52!}\}$ , where each  $\sigma_i$  is a distinct permutation.

Since each permutation in  $\Omega$  is equally likely when we shuffle a standard deck of cards, we have a uniform probability distribution.

The probability measure  $P$  assigns equal probabilities to all permutations in  $\Omega$ . Therefore, for any event  $A \subseteq \Omega$ , we have:

$$P(A) = \frac{|A|}{52!}$$

Specifically, for each individual permutation  $\sigma_i$  in  $\Omega$ :

$$P(\sigma_i) = \frac{1}{52!}$$

2. Probability of the each event is:

(a) **Finding the probability that the first two cards include at least one ace**

We can use the complement rule:

Probability (at least one ace in the first two cards) =

1 – Probability (no aces in the first two cards)

Now, let's calculate the probability of not getting any aces in the first two cards. There are 52 cards in the deck initially, and 48 of them are not aces. When we pick the first card, there are 48 non-ace cards out of 52 possibilities. After picking the first card, there are now 51 cards left in the deck, with 47 of them being non-ace cards. Therefore, the probability of not getting an

ace on the second card, given that the first card was not an ace, is:

$$\frac{48}{52} \cdot \frac{47}{51}$$

So, the probability that the first two cards contain at least one ace is:

$$1 - \left( \frac{48}{52} \cdot \frac{47}{51} \right) \approx 0.1494$$

**(b) Finding the probability that the first five cards include at least one ace**

Same as item (a) we can use the complement rule:

Probability (at least one ace in the first five cards) =

$$1 - \text{Probability (no aces in the first five cards)}$$

Same as previous item, the probability of not getting an ace on the first 5 cards, is:

$$\frac{48}{52} \cdot \frac{47}{51} \cdot \frac{46}{50} \cdot \frac{45}{49} \cdot \frac{44}{28}$$

So, the probability that the first five cards include at least one ace:

$$1 - \left( \frac{48}{52} \cdot \frac{47}{51} \cdot \frac{46}{50} \cdot \frac{45}{49} \cdot \frac{44}{28} \right) \approx 0.3147$$

**(c) The first two cards are a pair of the same rank**

**(d) The first five cards are all diamonds.**

There are 13 diamonds in a standard deck.

The probability of drawing a diamond on the first draw is  $\frac{13}{52}$  because there are 13 diamonds out of 52 cards.

Similarly, for the second, third, fourth, and fifth draws, the probabilities are as follows:

Second draw:  $\frac{12}{51}$

Third draw:  $\frac{11}{50}$

Fourth draw:  $\frac{10}{49}$

Fifth draw:  $\frac{9}{48}$

Because these are independent events, we can multiply these probabilities together:

$$\frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \cdot \frac{10}{49} \approx 0.00018184$$

So, the probability of drawing the first five cards as diamonds is approximately 0.00018184, or about 0.0182

(e) **The first two cards are a pair of the same rank**

## Problem 2

## Problem 3

## Problem 4

## Problem 5



## Problem 6

## Problem 7

For this problem a console program based on two libraries:

1. *requests* Python library → `DirectAPIClient.py`
2. *meteomatics.api* library → `HighLevelAPIClient.py`

### Retrieving Data Using DirectAPIClient Class

The `DirectAPIClient` class retrieves weather data from the Meteomatics API through the following steps:

1. It formats the time range for the query using the provided start date, end date, and interval.
2. The API URL is constructed by combining the base URL, time range, selected parameters, coordinates (latitude and longitude), and specifying the output format as JSON.
3. An HTTP GET request is sent to the constructed API URL, including your Meteomatics API username and password for authentication.
4. The response from the API is captured and checked for its status code.
5. In the case of a successful response (status code 200), the function parses the JSON response and extracts the weather data.
6. The weather data is processed to obtain time series for each selected parameter.
7. A Plotly figure is created, with traces (lines) added for each parameter to visualize trends over time.
8. The function includes error handling for errors based on the given HTTP status codes and provides informative error messages.
9. Also exception handling is implemented to catch any unexpected exceptions during the process.
10. The Plotly figure is displayed, allowing users to analyze and visualize the weather data.

## **How to Use The Program**

Run Python Main.py

Follow the prompts to select parameters, input dates and time intervals, and choose either the high-level or direct API client.

The program will then query the Meteomatics API, retrieve weather data, and display it as a Plotly graph.

You can visualize and analyze the trends of the selected weather parameters over a any time period that you want.

## Example

In this example, Precipitation, Wind speed, Pressure, Weather Symbol and UV trend over 24 hours is plotted.

```
"E:\EDU\Data Mining\dm-projects\hw1\repo\venv\Scripts\python.exe" "E:\EDU\Data Mining\dm-projects\hw1\repo\Main.py"
Enter 'list' to see available parameters or press Enter to continue:
Enter 'meteomatics' to use the high-level API client with the meteomatics library or enter 'requests' to use the requests module: meteomatics
Enter your Meteomatics API username: username@university.edu
Enter your Meteomatics API password: @P@ssw0rd
Enter latitude: 41.881833
Enter longitude: 12.496366
Enter weather parameters (comma-separated): 't_2m:C', 'precip_1h:mm', 'wind_speed_10m:ms', 'rel_pressure:hPa', 'weather_symbol:idx', 'uv:idx'
Enter start date and time (e.g., 2023-10-05 00:00:00): 2023-10-12
Enter end date and time (e.g., 2023-10-16 00:00:00): 2023-10-13
Enter data interval in hours: 1
Process finished with exit code 0
```

Figure 1: Console Output of the Example

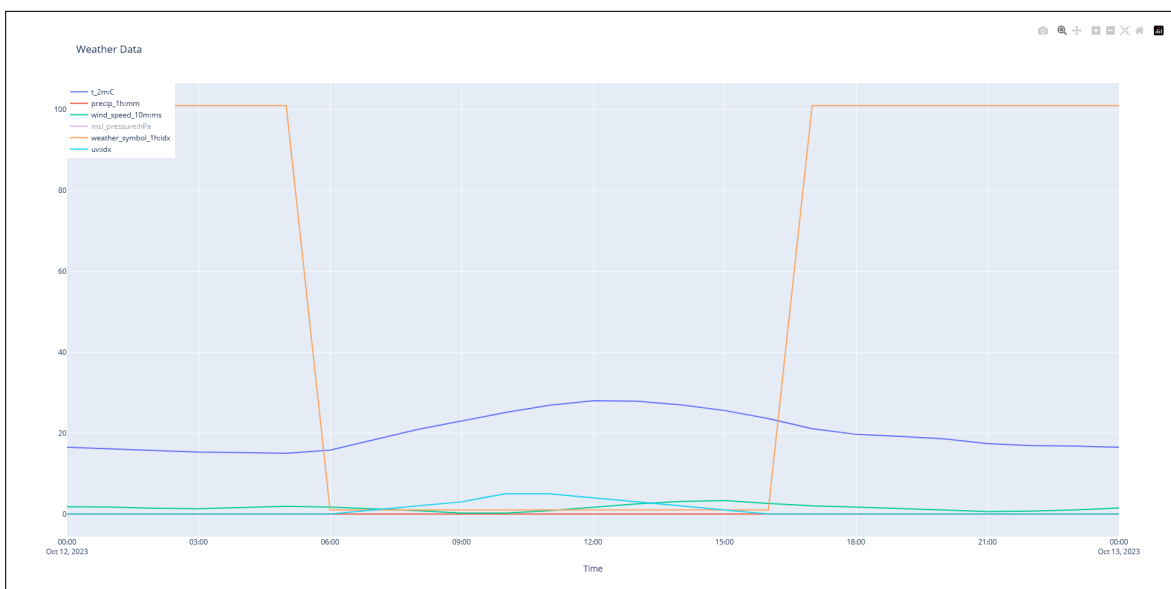


Figure 2: Plot of Precipitation, Wind speed, Pressure, Weather Symbol and UV trend over 24 hours