Mini-Project Instructions

Python for Data Analytics and Statistics
Metropolia University of Applied Sciences

2025

Project Overview

Each student must complete a mini-project in Python for Data Analytics & Statistics. You can choose to work individually or in a group of up to 3 students. If you work in a group, make sure all members agree on the project details and submit the form with identical information.
Complete the form to confirm your project details: Project Registration Form
Deadline: 9.4.2025 at 12:00 PM

Project Components

Components Breakdown

- **Load the Dataset and Explore Basic Information**
  - Open and inspect the dataset.
  - View the first few rows to understand its structure.
  - Check for missing data and data types.
- **Perform a Descriptive Summary**
  - Calculate basic metrics (mean, median, standard deviation).
  - Identify trends (e.g., average values).
  - Detect any unusual values or outliers.
- **Perform a Time-Based Analysis**
  - Group data by time (e.g., hourly, daily, weekly).
  - Identify trends, peaks, or seasonal patterns.
- **Visualize Data**
  - Create meaningful graphs and charts to present insights.
- **(Optional) Predictive Analysis**
  - If you missed more than 2 hands-on exercises, you can compensate by doing a predictive analysis.
  - Split the dataset into training and testing sets.
  - Use variables to predict outcomes (e.g., regression-based predictions).

Sessions Covered

Relevant Sessions

To successfully complete the mini-project, refer to the previous sessions:

- **Session 1 (5.3.2025)** - Python Basics, Data Types, File Handling.
- **Session 2 (13.3.2025)** - NumPy & Pandas, Data Cleaning, Time-Series Analysis.
- **Session 3 (19.3.2025)** - Descriptive Statistics, IQR, Outliers.
- **Session 4 (26.3.2025)** - Probability Distributions, Hypothesis Testing.
- **Session 5 (2.4.2025)** - Correlation, Regression Analysis.
- **Session 6 (9.4.2025)** - Matplotlib & Seaborn for Visualization.
- **Session 7 (16.4.2025)** - Mini-Project Walkthrough.

Submission Guidelines

Submission Instructions

- Upload your Python script (.py) or Jupyter Notebook (.ipynb) to GitHub.
- Submit your GitHub link via the project form.
- Submit your project at: Metropolia Submission Link.
- **Deadline:** 22.4.2025 at 12:00 PM.

Final Presentation (23.4.2025)

Presentation Details

- **Time Limit:** 5-10 minutes.
- Present your code or prepare a short slide deck (5-7 slides).

GitHub Requirement

GitHub Instructions

- If you don't have a GitHub account, create one immediately.
- Watch this tutorial on how to set up GitHub: How to Create a GitHub Repository.

Contact Instructor

Instructor Contact

Hamed Ahmadinia, Ph.D.
Instructor – Python for Data Analytics & Statistics
Email: hamed.ahmadinia@metropolia.fi