Syllabus: Introduction to Social Data Science

Course Information

- **Course Code:** ULI2008.1 – Sosyal Veri Bilimine Giriş / PSIR3066 – Introduction to Data Science
- **Schedule:** Monday, 3 hours (14:00–17:00)
- **Classroom:** Faculty of Political Science, Room RTE.S1.124
- **Office Hours:** Monday, 09:00–11:00
- **Office Location:** Faculty of Political Science, Room RTE.S1.131
- **Credits & ECTS:** 3 credits, ECTS: 5
- **Course Type:** Elective
- **Instructor:** Assoc. Prof. Hakan Mehmetcik
- **Contact:** hakan.mehmetcik@marmara.edu.tr

Course Objectives

This course provides an introduction to **R**, a free and open-source statistical computing environment and programming language widely used in **data analysis, statistical research, and industrial applications**. The primary goal of this course is to equip students with the ability to use **R for problem-solving and practical applications** in social data science. Students will gain hands-on experience with data manipulation, visualization, statistical computation, and analysis beyond basic spreadsheet tools such as Excel.

R is a versatile programming language that enables efficient **data processing, statistical modeling, and visualization**. It runs on multiple platforms, including **UNIX, Linux, Windows, and macOS**.

This course is designed for students who want to learn the fundamentals of R and its applications in social data science. The content focuses on:

- The **basics of the R environment** and how to navigate it effectively.
- **Arithmetic and logical operators**, data structures, variables, and data types in R.

- **Data manipulation and transformation** techniques, including working with R packages for visualization, reporting, and statistical analysis.
- **Importing, processing, and transforming datasets** for analytical purposes.
- Core **data science concepts**, including **parametric and non-parametric statistical tests**.

By the end of the course, students will have a strong foundation in **R programming** and its **applications in data science**, enabling them to analyze and visualize real-world datasets effectively.

Course Learning Outcomes

By the end of this course, students will be able to:

- **Use R and RStudio effectively** for data analysis and programming.
- **Process and analyze data** using R's built-in functions and libraries.
- **Create visualizations and reports** to communicate data-driven insights.
- **Apply statistical methods** for data exploration and hypothesis testing.

These skills will provide students with a strong foundation in **data analysis and statistical computing**, preparing them for **more advanced data science and problem-solving tasks**.

Course Readings

Primary Textbooks (Required Readings)

Modern Data Science with R

R for Data Science

Supplementary Resources (Optional Readings & References)

- An Introduction to R (pdf) by W. N. Venables, D. M. Smith, and the R Development Core Team
- Advanced R (2nd Ed) by Hadley Wickham (advanced)
- Intro to R video lectures by Google Developers
- R Programming wikibook
- Using R for Data Analysis and Graphics by J. H. Maindonald
- The R Inferno by Patrick Burns (advanced)

Required Materials for the Course

Hardware Requirement

- Students **must bring a laptop** to class for hands-on exercises.

Software Requirement

To participate in this course, students need to install **R and RStudio** on their personal computers.

1. **Download and Install R**
    - First, download and install **R** from The R Project.
2. **Download and Install RStudio**
    - After installing R, download and install **RStudio** from Posit.
3. **Install Required R Packages**
    - During the course, students will need to install and use the **tidyverse** and **RMarkdown** packages. Instructions for installing these and other required packages will be provided in class.

Getting Help

If you need assistance, consider the following steps:

1. **Self-Help Resources**
    - Use **R Help (?command)** within RStudio.
    - Search for solutions on **Google** or **AI platforms** (e.g., **ChatGPT, DeepSeek**).
2. **Ask Questions in Google Classroom**
    - Post your questions in the **Google Classroom discussion board** for peer and instructor support.
    - If you know the answer to a classmate's question, feel free to help!
3. **Schedule a Virtual Meeting**
    - If additional support is needed, students can **schedule a virtual meeting** via **Google Meet** with the instructor.

Attendance and Participation Requirements

Research indicates that **regular attendance** is one of the strongest predictors of academic success.

According to the university's **Education and Teaching Regulations**, students are required to:

- Attend at least **70% of theoretical classes**.
- Attend at least **80% of practical sessions** (if applicable).

Please consider these attendance requirements carefully before enrolling in this course!

Grading & Final Project Requirements

Grading Breakdown

- **Midterm Project – Phase 1:** 40%
- **Final Project – Phase 2:** 50%
- **Participation:** 10%

For exact deadlines, please check the university's online system (OBYS).

Each student is required to conduct a **data analysis using R** based on a dataset sourced from **Data is Plural**. The project consists of two phases:

Midterm Project – Phase 1: *(Problem Definition & Data Collection)*

1. **Problem Definition & Data Selection:**
   - Identify a **social issue** you want to analyze.
   - Formulate a **clear research question** and determine the relevant dataset.
   - Data **must be selected from** Data is Plural's dataset archive.
2. **Data Cleaning & Exploratory Data Analysis (EDA):**
   - Prepare the dataset for analysis by handling **missing values, outliers, and irrelevant data**.
   - Conduct **Exploratory Data Analysis (EDA)** to identify patterns and relationships in the data.

Final Project – Phase 2: ***(Data Modeling & Reporting)***

3. **Data Modeling:**
    - Based on EDA insights, select **appropriate statistical or machine learning models** to answer your research question.
    - Justify your model choice and interpret the results.
4. **Presentation & Report Writing:**
    - Clearly present your findings through **visualizations, a written report, or a presentation**.
    - The report should include all project phases, from **problem definition to conclusions**.

For exact submission deadlines, please check Marmara University's BYS system.

Academic Integrity & Ethical Guidelines

To ensure a **fair, transparent, and ethical learning environment**, students must adhere to the following principles:

1. **Academic Honesty** – Plagiarism and cheating are serious offenses and will not be tolerated. Any violations may result in disciplinary action.
2. **Exam Conduct** – Students must follow all exam rules and refrain from using unauthorized materials (e.g., books, notes, electronic devices).
3. **Original Work** – All assignments and exams must reflect the student's own work. Proper citation is required for any external sources.
4. **Respect for Peers** – Students should engage respectfully with classmates and avoid unauthorized collaboration on assignments.
5. **Instructor Communication** – If you have concerns about academic integrity or ethical dilemmas, discuss them openly with the instructor.
6. **Accountability** – Students are responsible for their own learning, participation, and adherence to ethical guidelines.
7. **Consequences of Violations** – Breaching these principles may result in academic penalties, such as assignment failure or disciplinary actions by the university.

Maintaining **integrity and fairness** in academics fosters a positive learning environment and supports students' intellectual growth.

Accessing Google Classroom

To join the course on **Google Classroom**, follow these steps:

1. **Sign in to your Google Account** and use the invitation link:
   Join Google Classroom
2. Alternatively, you can manually join:
   - Go to Google Classroom.
   - Sign in with your **Google Account**.
   - On the "My Classes" page, click **"Join"**.
   - Enter the **class code:** `xweaxw6` and click **"Join"**.

Important Notes:

- Ensure that you enter the **correct class code** to avoid errors.
- Once you join, you will have access to **assignments, announcements, and course materials**.
- You can also access Google Classroom from your **mobile device**.
- Make sure you are using the **latest version** of the Google Classroom app for a smooth experience.

Week	Date	Type	Topic	Readings

Week	Date	Type	Topic	Readings
1	February 17, 2025	Lesson	Introduction to R, RStudio, and R Markdown	MDSR Appendix A-B-C, MDSR 1, R4DS 1
2	February 24, 2025	Lesson	Data Visualization with ggplot	MDSR 2, R4DS 3
3	March 03, 2025	Lesson	Data Visualization with ggplot2	MDSR 3, R4DS 3
4	March 10, 2025	Lesson	Data Frames and Data Manipulation	MDSR 4-5, R4DS 5-8
5	March 17, 2025	Lesson	Data Summarization and Pipes	MDSR 6-7, R4DS 9-16
6	March 24, 2025	Lesson	Data Communication, Data Import and Workflows	MDSR Appendix D, R4DS 26
7	April 07, 2025	Lesson	Data Communication, Data Import and Workflows	MDSR Appendix D, R4DS 267
8	April 21, 2025	Lesson	Linear Regression, Inference and Modeling	MDSR Appendix E, R4DS 24
9	April 28, 2025	Lesson	Descriptive Statistics, Inference and Modeling	MDSR 9-10, R4DS 22-23
10	May 05, 2025	Exam	NA	NA
11	May 12, 2025	Lesson	Machine Learning, Modeling	MDSR 10-11, R4DS 25
12	May 19, 2025	Lesson	Modeling, Text as Data	MDSR 19, R4DS 14
13	May 26, 2025	Lesson	Project Sample	Hands-out
14	June 02, 2025	Lesson	Project Sample	Hands-out
15	June 16, 2025	Lesson	Project Sample	Hands-out
16	June 23, 2025	Lesson	Review	NA
17	June 30, 2025	Exam	NA	NA

Additional Information

- The course **content and structure** may be modified at the instructor's discretion. Any changes will be communicated in advance.
- Students are encouraged to reach out to the **instructor or teaching assistant** if they have any questions or concerns regarding the course. Contact details are provided above.
- **Note:** This syllabus outlines the **general structure and teaching methods** of the course. Additional course materials and a detailed schedule will be provided as the course progresses.

Wishing you a successful semester!