

Text as Data with R

This course is designed to provide participants with comprehensive skills for working with text data using R. The participants first will learn how to import and preprocess textual data, ensuring that raw data is transformed into a clean format suitable for analysis. The course will then introduce widely used practices such as tokenization and text wrangling, allowing learners to break down and manage the data efficiently. As participants build their foundational skills, they will engage in exploratory analyses—such as frequency analyses and the creation of word clouds—that help reveal underlying patterns in the text. This exploratory phase is crucial for gaining initial insights into the data’s structure and content. Following this, the course covers fundamental sentiment analysis and text classification techniques, equipping participants with the tools to determine the emotional tone and categorize the content of the text. Lastly, the participants will explore simple topic modelling and clustering techniques, which enable them to identify recurring themes and group similar texts together. The course concludes with an emphasis on visualization and interpretation, teaching students to present their findings clearly and meaningfully to inform decision-making. Through these interconnected modules, the course ensures that participants are proficient in technical methods and can translate their analytical results into actionable insights.

This course will benefit junior scholars, and students new to text analytics and programming in R will find the course accessible and practical for building foundational skills in data import, preprocessing, and fundamental text analysis. Moreover, it would also be helpful for experienced scholars looking to integrate new methodologies into their research to appreciate the advanced topics, such as sentiment analysis, text classification, and topic modelling, which complement their existing analytical toolkits.

Prerequisites:

Familiarity and basic knowledge of R and R studio



Emre Toros

Emre Toros is a professor at İhsan Doğramacı Bilkent University, Department of Communication Design. He previously held academic leadership roles at Hacettepe University and Atılım University, where he served as a professor dean and department chair. His international experience includes visiting positions at Harvard University, King’s College London, and Stanford University, supported by prestigious fellowships such as Fulbright and the British Academy. His research focuses on political behavior, public opinion, electoral integrity, and advanced research methodologies, including text mining and social media analysis. He has led multiple research projects funded by national and international institutions, exploring topics such as electoral trust, campaign strategies, and climate resilience. Professor Toros has published extensively in leading academic journals and contributed to books on political communication and electoral studies. His expertise in quantitative methods, survey experiments, and data visualization informs his research and teaching, spanning courses on research methods, statistics, political behavior, and social media analysis.