

# SHAH ALI GARDEZI

[Gardezi\\_Syed-Shah@student.ceu.edu](mailto:Gardezi_Syed-Shah@student.ceu.edu) • +36 20 335 6475 • [www.linkedin.com/in/shah-ali-dataanalyst](https://www.linkedin.com/in/shah-ali-dataanalyst)

## ACAMEDIC PROJECTS

---

### Course: Data Engineering with MySQL

- Project 1: Created an *operational data layer* in MySQL from the relational dataset (Formula One racing). Then designed an *analytical layer* – a denormalized data structure using the operational layer to create a table in MySQL. Constructed an *ETL* pipeline using Triggers and Stored Procedures and created Views as *Data Marts*. Project report can be found on my GitHub [here](#)
- Project 2 (Group Project): Created a KNIME-based workflow using 3 data sets, 2 of which were accessed through online APIs. Constructed an ETL pipeline in MySQL to render a datatable. Using this datatable, analysis was performed and visualization were generated. Project report can be found on my GitHub [here](#)

### Course: Data Analysis 2: Finding patterns with Regression

- Project deliverable included finding the association of a *dependent variable* ( $y$ ) with the *explanatory variable* ( $x$ ) including *confounding variables* ( $z$ ) based on the cross-sectional dataset on supermarket sales. After performing Exploratory Data Analysis (EDA) several OLS regression models were created each included different categorical and numerical variables. The association of Loaves of bread ( $y$ ) with price per unit of bread ( $x$ ) was evaluated. Project report can be found on my GitHub [here](#)

### Course: Data Analysis 3: Prediction & Machine Learning Analysis

- Build several price prediction models using Airbnb's cross-sectional dataset for the city of Puglia, Italy. Evaluated the performance of each prediction model and made recommendation of the best models based on RMSE and BIC values. Project report can be found on my GitHub [here](#)

### Course: Data Engineering 2: Cloud Computing

- Conducted a sentiment analysis comparing 2 newspaper articles on a similar news using Amazon Web Services (AWS) tool; *Amazon Comprehend*. Project report can be found on my GitHub [here](#)

### Course: Web Scrapping using R

- Scraped all of information presented on [CSIMarkets](#) website about stock market of 500 companies. Arranged them alphabetically. The information was then to be added in a dataframe which was then stored into and RDS. Visualizations were plotted of the some parameters from the created the dataframe as well. Project report can be found on my GitHub [here](#)

### Course: Data Visualization in Tableau

- Created 2 page interactive dashboard using Human Resource dataset provided in the course to answer a number of analytical questions about the company resources. Visualization included bargraphs, donut charts, geo-maps and area-under-the-graph charts. Incorporated filters, drill down and visual tooltips. Project report can be found on my GitHub [here](#)

### Course: Practical Data Visualization in R

- Used built-in R themes such as *plotly* and *gganimate* using *data.table* package to answer the analytical questions concerning the dataset on deforestation. Project report can be found on my GitHub [here](#)