

# ZEYNEP TUGCE SAHAN

tugcesahan@gmail.com | linkedin.com/in/tugcesahan | tsahan.github.io | (765) 205-1262

## SUMMARY

- Committed, self-motivated, and proactive Data Engineer with more than 6 years of experience in software development and 4+ years of experience working with data, including but not limited to data analytics & visualization, ETL, data migration, and integrating third-party systems
- Hands-on experience in agile environment (Kanban and biweekly sprints) and Git-based source control gained while working at start-ups, academia, IT, and government
- Strong communication skills and ability to engage with all levels of an organization

## EDUCATION

### Master of Science (MS)

December 2020

Purdue University / West Lafayette, IN

Major: Industrial Engineering | CGPA: 3.97

### Bachelor of Science (BS)

June 2017

Antalya Bilim University / Antalya, Turkey

Majors: Computer Engineering, Industrial Engineering

## PROFESSIONAL EXPERIENCE

### Data Engineer

February 2021 - Present

enFocus, Inc / South Bend, IN

- Build a data pipeline to monitor states of scheduled tasks and report results using Python, Sharepoint, Power Automate, and Azure DevOps. (<https://tsahan.github.io/project/scheduled-tasks-monitoring-in-devops/>)
- Manage stored procedures that pull in data from linked servers, make comparisons, and update the data in the local server.
- Transform data from SQL Server and transfer it to an external application by making API requests using Python pyodbc, requests, pandas, and json.
- Develop a custom Python module that tracks script failures, enables logging, and sends details in an email.
- Create SSIS packages that utilize SQL and Python to identify voided invoices, increase collection rates, and save over \$50K annually.
- Work closely with a cross-functional team to troubleshoot issues and maintain automated scripts in MS SQL Server.
- Leverage Git and Git repositories for version control, code reviews, and maintaining documentation.

### Business Intelligence Developer Intern

May 2020 - August 2020

enFocus, Inc / South Bend, IN

- Automated processes using SSIS, Python, T-SQL, and SQL Server to eliminate ~18 hours of manual work per week.
- Prepared interactive dashboards using MS Power BI to visualize data for technical and non-technical audiences.
- Designed custom reports that update with user input in SAP Crystal Reports to save ~4 person-hours per month.

### Graduate Research Assistant

August 2017 - May 2020

Purdue University / West Lafayette, IN

Homeowner-Level Decision Support System for Mitigating Coastal Flood Risk in Louisiana

- Developed a decision support tool to help homeowners make better decisions about managing flood risks to their properties using JavaScript, HTML, CSS, ArcGIS API, and Google Maps API.
- Combined data from SQL Server and ArcGIS Image Server; and analyzed it using Python.
- Communicated complex quantitative analysis in a clear, precise, and actionable manner.
- Collaborated with CPRA and USGS to make the product available to 2.3 million coastal Louisiana residents. (<https://tsahan.github.io/project/flood-risk-dss/>)

### Data Analyst Intern

June 2019 - August 2019

Purdue University Data Analytics and Information Office / West Lafayette, IN

- Designed a relational database to combine data from multiple data sources.
- Wrote queries for extracting useful information and providing insights.
- Created a user interface for data updates and anomaly detection using VBA.

## PROJECTS

### Text Classification using NLP Algorithms

October 2021 - Present

- Prepare data for training and testing using Python packages such as pandas, sklearn, numpy, and nltk.
- Classify highly complex text data using natural language processing algorithms like Naive Bayes, SVM, SGD, XGBoost, and BERT to eliminate ~2 hours of manual work per week.
- Grid search over pipeline to optimize parameters for entire workflow.

## HR Analytics - Employee Turnover Prediction

August 2018 - October 2018

- Analyzed data using descriptive, predictive, and prescriptive analytics methods in R.
- Created machine learning models using logistic regression, decision tree, SVM, and XGBoost to predict the probability of employees leaving a company based on their characteristics with 0.9 AUC.
- Developed a dashboard to present exploratory data analysis results, make predictions with user input, and provide actionable insights using R packages like shiny, ggplot2, caret, tidyr, and dplyr.  
(Available at <https://tugcesahan.shinyapps.io/HR-Analytics/>)

## Airlines Revenue Management Decision Support Tool

April 2017 - May 2017

- Created an LP model to maximize profit by finding the number of seats allocated for cabin classes in each flight.
- Performed a sensitivity analysis under different scenarios using Excel What-If Analysis tool.
- Utilized advanced level abilities including vlookups, index, match, pivot tables, and Macros in Excel to provide an interactive user interface.

## TECHNICAL SKILLS

**Programming Languages, Frameworks, and Databases:** Python, R, MATLAB, VBA, C, SQL, MS SQL Server, NoSQL, Javascript, JupyterLab, Git, MapReduce, Spark, Airflow, Redshift

**Applications and Platforms:** Power BI, Tableau, MS Power Automate, Arena Simulation, MS Visual Studio, SSMS, SSIS, ArcGIS, SAP Crystal Reports, Azure DevOps, Postman

## CERTIFICATIONS

- |  |   |
|--|---|
| <ul style="list-style-type: none"><li>• Google IT Automation with Python Professional Certificate<ul style="list-style-type: none"><li>- Crash Course on Python</li><li>- Using Python to Interact with the Operating System</li><li>- Introduction to Git and GitHub</li><li>- Troubleshooting and Debugging Techniques</li><li>- Configuration Management and the Cloud</li><li>- Automating Real-World Tasks with Python</li></ul></li><li>• Getting Started with Power BI Desktop - Coursera</li><li>• AWS Certified Cloud Practitioner - A Cloud Guru</li></ul> | <p>November 2020</p> <p>August 2020<br/>July 2020</p> |
|--|---|