Somang (Elena) Song

elenassong@alumni.unc.edu | (919) 903-7941| LinkedIn |Cary, NC

EXPERIENCE

University of North Carolina at Chapel Hill — **Chapel Hill, NC**
Data Analyst Intern — May – Sept. 2022

- Developed Python libraries to characterize and analyze diagnostic features related to social risk factors, enhancing the capability to understand health disparities
- Designed and executed statistical models, such as linear regression and two-way ANOVA, to investigate the impact of social determinants on patient care quality
- Utilized Tableau to create data visualizations, illustrating the impact of socio-economic factors on healthcare outcomes based on 22k+ EPIC Electronic Health Records (EHRs) from NC's Women's Hospital

Chungnam National University Hospital — **Daejeon, South Korea**
Research Associate — Sept. 2020 – June 2021

- Assisted in optimizing the SQL database performance to ensure efficient and timely access to critical patient data, reducing query response times by 25%
- Designed and implemented a comprehensive COVID-19 patient outcome reporting system, providing daily, accurate, and real-time insights for clinical evaluations during the pandemic
- Transformed PubMed XML data into a structured CSV format for SQL database ingestion and executed queries to analyze medical research

SELECTED PROJECTS

Sports Statistics and Predictions (Python, SQL)
Developed a machine learning model for predicting sports match outcomes by collecting the European Soccer match database. Conducted in-depth exploratory data analysis (EDA) to identify relevant patterns and relationships. Leveraged a diverse array of machine learning algorithms, including logistic regression and XGBoost to create a robust prediction model. Evaluated model performance using metrics like accuracy, precision, and recall with cross-validation to ensure reliability.

Cross-Platform Design for Subscriptions – Database System Project (MySQL, PHP)
Architected Enhanced Entity-Relationship (EER) diagram to establish a data model with normalized relations and functional dependencies tailored to bicycle subscription services. Utilized PHP and MySQL in the XAMPP environment to facilitate efficient data loading such as time slot availability and reservation features.

System Analysis Project (Collaborated with The Center for Discovery, New York)
Researched electronic health record systems such as SigmaCare and eClinicalWorks. Interviewed healthcare professionals to analyze operational issues and assess the advantages and disadvantages of hybrid EHR. Utilized Agile methodology, Sequence Model, and Affinity Diagram for effective system analysis, focusing on optimizing eMAR benefits.

TECHNICAL CAPABILITIES

- **Programming Languages:** MySQL, Python
- **Development Tools:** Git and GitHub (Version Control), CI/CD, Linux and Shell Scripting (OS), AWS (Cloud)
- **Data Visualization:** Matplotlib, Seaborn, Tableau
- **Skills:** Data Analysis, Data Modeling, Machine Learning, Statistical Testing, Project Management (Six-Sigma, Agile)

EDUCATION

University of North Carolina at Chapel Hill — **Chapel Hill, NC**
MPS in School of Information and Library Science (CHIP) — 2022
Relevant coursework: Database Systems 1/2, System Analysis, Programming for Data Analysis, Statistic Analysis

Kyung-Hee University — **Seoul, South Korea**
M.S in Biotechnology (GPA: 3.77/4.0) — 2019
B.S in Genetic Engineering (GPA. 3.20/4.0) — 2017