DUKE (SUNG HOON) YANG

shy256@nyu.edu | sunghoonyang.github.io | +1 (929) 319-8663

Education

NEW YORK UNIVERSITY   NY, UNITED STATES
M.S. in Applied Urban Science & Informatics   **To graduate in Aug '19**

UNIVERSITY OF CALIFORNIA, BERKELEY   CA, UNITED STATES
B.A. in Economics   **Graduated in Dec '15**

Employment History

Feb 2016 – Aug 2018

Data Engineer   **COUPANG** Seoul, South Korea

- Migrated on-premise Data platform (IBM PureData) to AWS. Spearheaded the DQ project, with automated reporting. Released and administered Redshift data warehouses.
- Built and Owned a Flask application that automates scheduled job generation using Jinja, SqlAlchemy, and Airflow.
- Designed TMS domain's mart tables to automate weekly business reports.
- Designed and maintained Airflow DAGs & Operators for various source databases, including unsupported MongoDB or Greenhouse API.
- Built ETL SLA monitoring tool using node.js with express & D3.js, serviced via email and Slack.

May 2015 – Aug 2015

Data Analyst   **IDEASCALE** Berkeley, CA

- Administered Google Analytics(GA) by building custom metrics to track customer behavior.
- Designed business reports in GA dashboard.

Jan 2014 – Apr 2014

Data Analyst   **NATIONAL ASSEMBLY, REPUBLIC OF KOREA** Seoul, South Korea

- Owned data collection from various sources (PDF invoices, Excel, and API data), and produced analytic reports for MP Min Seok Ahn to be used in public hearings.
- Published reports of historical budget spending per category, supporting ad hoc drill-down on user-defined dimensions.
- Enabled automatic reporting using VBA.

Jan 2012 – Nov 2013

Tank Driver, Squad Leader   **REPUBLIC OF KOREA ARMY** Ga-Pyeong, South Korea

- Operated and maintained a K-21 armored vehicle.
- Taught English for 16 months.

Skills

Technology Skills: Python, SQL, NoSQL, Bash, Javascript, AWS(S3, EC2, EMR, Redshift), Distributed System(Hadoop, Spark, Zeppelin), Unit Testing, Agile Method(git, Jira, Sprint, etc.), Database Administration, MVC App Development (Ruby on Rails, Flask, node.js), ArcGIS
Data Skills: Modeling(Dimension & Fact Design), Quality Assurance, Analysis(BI, ETL, KPI Aggregation), Visualization(Tableau, PowerBI, D3.js, matplotlib, etc.), Machine Learning, Spatial Analytics (Clustering, Autocorrelation, Spatial Interpolation & Regression)
Hobbies: Logic Pro, Music Composition, SuperCollider