

# Ankit Dhall

Email : ankit.dhall@outlook.com

Mobile : +1 518-620-7401

[linkedin.com/in/ankit-dhall](https://www.linkedin.com/in/ankit-dhall)

[github.com/dhallankit](https://github.com/dhallankit)

## EDUCATION

---

### University of Maryland

Master of Information Management (STEM); GPA: 3.87/4

College Park, MD

August 2018 - May 2020

### University of Mumbai

Bachelor of Engineering; Major - Electronics and Telecommunication Engineering

Mumbai, MH, India

August 2012 - June 2016

## SKILLS

---

**Programming:** Python, SQL, Shell

**Databases:** MySQL, Postgres, BigQuery

**Big Data:** Hadoop, PySpark

**Platform:** Unix, Linux, Mac, Windows

**Virtualization:** Docker, Kubernetes

**Technology:** GraphQL, Apache Airflow, FastAPI, Looker

**Cloud:** Google Cloud Platform (GCP), Azure

**Project Management:** Jira, Confluence, Git, Jenkins CI/CD

## EXPERIENCE

---

### Forbes

Data Engineer, BI

Jersey City, NJ

Apr 2021 - Present

- Owned and maintained Cloud Composer(Airflow) DAGs for the BI team's daily loader process
- Created Python based data enrichment process to decrease the processing time from hours to minutes
- Enhanced daily alerts system with better monitoring using BigQuery API and performed troubleshooting fixes

### PwC

Data Engineer Associate

Tampa, FL

Jul 2020 - Apr 2021

- Translated Docker based Airflow and FastAPI apps to Kubernetes on Azure to increase scalability and performance
- Setup GraphQL engine on Kubernetes and integrated with existing applications increasing performance by 3x
- Troubleshot bugs and errors in all environments for Airflow and FastAPI applications on Kubernetes and Docker
- Created MySQL DB migration script with parallel processing resulting in faster and reliable concurrent migration
- Increased flexibility of applications to make more dynamic using configuration scripts and Vault integration

### University of Maryland

Faculty Data System Assistant

College Park, MD

Aug 2019 - May 2020

- Implemented Python ETL process for data analysis of Research Awards of 1500+ faculty with version control on Git
- Performed extensive data cleaning and transformations using Python and Google App scripting for 50000 records
- Conducted training and pilot testing to generate user insights to understand desired UI and functional changes

### PwC

Data Engineer Intern

Tampa, FL

Jun 2019 - Aug 2019

- Created ephemeral Docker container-based clusters using Python to reduce Azure server consumption costs
- Automated and enhanced data flow on Swagger based REST APIs for CAMS (Common Asynchronous Management System) platform using Airflow to manage track databases, clusters, images and engagement teams
- Unified ERP databases for XDM (Cross Data Model) to build a uniform data model for engagement teams
- Owned module in QA testing effort providing on time delivery and ensuring robustness of product

### Reliance Jio Infocomm Ltd

Assistant Manager - Integration Technology

Mumbai, MH, India

Sep 2016 - Jun 2018

- Designed and analyzed reliable data structures & schemas associated with business requirements to unify platforms
- Created middleware workflows for orchestrated events to onboard customers using Enterprise Architecture Integration

## PROJECTS

---

### COVID19 Sentiment Analysis on Twitter using Big Data

- Implemented Pub-Sub model to collect tweets around COVID19 based on selected hashtags for over a period of 9 days
- Used Apache Beam based Cloud Dataflow to gather tweets (2.8 GB) from Pub-Sub and store in BigQuery table
- Applied Sentiment Analysis to effectively compare between News Sources, Official Medical Orgs and other users