DANG NHAT

DATA ENGINEER

12/04/1995

Go Vap District, HCMC

+84 0354883673

dangnhatsimon@gmail.com

Profile

LinkedIn

Skills

- Basic:
  - Data Structures
  - Algorithms
  - Prob & Stats
- Programming:
  - Python, R
  - Java, Scala
  - JavaScript, LaTeX
- Scripting:
  - Bash, Shell
- Database:
  - SQL, NoSQL
- Big Data:
  - Spark, Kafka
- Cloud:
  - GCP, IBM
- Data Visualization:
  - Tableau, Power BI

Detail-oriented | Structured thinking | Analytical thinking | Hard working | Teamwork | Data Storytelling

Working Experience

Oct 2022 – Present

Data Engineer PECC2

Contribute to Big Data Development Plan for Company. Hadoop, Spark for Data Analytics. Running automation with Python, Bash scripting on Ubuntu. Version control system like Git. Monitoring health of server computers.
Getting data like power output, weather, current and voltage from site, through pipelines to database server at company, after processing and manipulating, data will be analyzed, visualized in apps. All reports and dynamic dashboard will be sent streaming to the end users to help them improve experiences, making decisions.
Core technologies: Big Data Analytics, ETL, Data Engineering, Python, Bash, Git, Ubuntu.

Oct 2019 – Oct 2022

Mechanical Piping Engineer PECC2

BIM Manager and Mechanical Lead. Thermal-Mechanical and piping engineering.

Project

Sep 2023 – Present

IBM Data Engineering **Data Engineer**

Implement webscraping and use APIs to extract data with Python.
Create, design, & manage relational databases & apply database administration (DBA) concepts to RDBMSs such as MySQL, PostgreSQL, & IBM Db2. Create views, transactions, stored procedures and joins.
Develop working knowledge of NoSQL & Big Data using MongoDB, Cassandra, Cloudant, Hadoop, Apache Spark, Spark SQL, Spark ML, and Spark Streaming.
Implement ETL & Data Pipelines with Bash, Airflow & Kafka; architect, populate, deploy Data Warehouses; create BI reports & interactive dashboards.
Core technologies: Data Science, ETL & Data Pipelines, RDBMS, NoSQL and Big Data, Apache Spark, Python Programming, Data Analysis, SQL.

Aug 2023 – Aug 2023

Automation **Data Engineer**

Scale and convert images using PIL.
Automate updating catalog information. Fetching and working with supplier data images. Process Text Files with Python Dictionaries. Uploading images, descriptions to web server. Write a script that summarizes and processes sales data into different categories. Generate a PDF report and send it through email using Python. Write a script to check the health status of the system.
Core technologies: Bash, Linux, Python.

Aug 2023 – Aug 2023

Data Engineering **Data Engineer**

Ingest data from external sources (SAS,STATA,HDF5,MAT,..), manipulate, process data.
Database Design.
Use PySpark to Create a Data Transformation Pipeline.Understand Resilent Distributed Datasets (RDDs) of Spark, using pyspark to store data in distributed file system, transformation and actions datasets.
Spark SQL for Big Data Analytics. Create, import, inspecting, cleaning, querying and saving data with pyspark.
Testing data pipeline.
Manage and Orchestrate Workflows with Apache Airflow, DAG schedule.
Core technologies: Python, Spark, ETL, Airflow, SQL.

DANG NHAT

DATA ENGINEER

About Me

I am a results-driven Data Engineer with a passion for creating efficient data pipelines and optimizing data workflows. My expertise lies in designing and implementing robust data warehouses, modeling data structures for optimal performance, and integrating diverse data sources. With hands-on experience in data transformation and real-time processing, I am adept at turning raw data into actionable insights.

Languages

- Vietnamese
- English
- Russian
- Chinese

Project

Jul 2023 – Jul 2023 **freeCodeCamp** — **Data Analyst**

Mean-Variance-Standard Deviation Calculator: output the mean, variance, standard deviation, max, min, and sum of the rows, columns, and elements in matrix.
Demographic Data Analyzer: analyze, manipulate data from the 1994 Census database.
Medical Data Visualizer: converting, cleaning data, explore the relationship between cardiac disease, body measurements, blood markers, and lifestyle choices.
Page View Time Series Visualizer: analyze and visualize a dataset containing the number of page views each day on the webpage, understand the patterns in visits and identify yearly and monthly growth.
Sea Level Predictor: analyze the global average sea level change since 1880, predict the sea level change through year 2050.
Core technologies: Data Analysis with Python by using library Numpy, Pandas, Seaborn, Matplotlib.

Oct 2022 – Oct 2022 **Cyclistic Bike Share** — **Data Analyst**

Combine, cleaning, manipulate, analyze and visualize to get insights from 1 millions data casual and annual membership riders. Identify trend from Cyclistic's historical bike trip data and then planning digital media could affect marketing tactics.
Core technologies: MySQL, Python, R, Tableau.

Oct 2022 – Oct 2022 **FitBit Fitness Tracker** — **Data Analyst**

Combine, cleaning, manipulate, analyze the user data from FitBit Fitness Tracker to gain insights into how consumers are using the FitBit app and discover trends for Bellabeat marketing strategy.
Core technologies: MySQL, Python, R, Tableau.

Education

Undergraduate Study

2014 – 2019 **B.Sc. in Nuclear Power Engineering and Thermal Physics** — Russia

Title: Analysis of the transient mode in the VVER-1000 core when two main circulation pumps are switched off at rated power.
Supervisors: Prof. Leskin Sergei Terent'evich.
Knowlegde: Nuclear Power system and Nuclear Reactor, especially WWER-1000, Steam Generator and Heat Exchangers, Steam Turbine, Industrial Pump, Nuclear reactor safety systems, Valves
GPA: 4.6/5.0

Certificate

Sep 2023	Google IT Automation with Python	Google
Sep 2023	Problem Solving (Basic) Certificate	HackerRank
Aug 2023	Building Data Engineering Pipelines in Python	DataCamp
Aug 2023	Data Analysis with Python	freeCodeCamp
Aug 2023	Python Intermediate	Sololearn
Aug 2023	SQL Intermediate	Sololearn
Jul 2023	Big Data 101	IBM
Jul 2023	Hadoop 101	IBM
Jul 2023	Python (Basic) Certificate	HackerRank
Jul 2023	SQL (Intermediate) Certificate	HackerRank
Jan 2023	Google Data Analytics	Google