

# LUKE CHUGH

Phone: +44 7780012032 | Email: [lukechughformal@gmail.com](mailto:lukechughformal@gmail.com)

LinkedIn: <https://www.linkedin.com/in/luke-chugh-2b2043181/>

Portfolio Website: <https://lukechughformal.wixsite.com/lukechughportfolio>

I am a 5 star Advance SQL coder on HackerRank and have solved 500+ LeetCode questions. I've implemented end to end projects in data science, converted them into WebApps and deployed them on cloud. Currently seeking an opportunity as a data analyst or a data scientist so that I can utilize the learnt skills to solve business problems and help the organisation grow.

## EDUCATION

---

**M.Sc. Data Science - Northumbria University, Newcastle Upon Tyne, United Kingdom**

Duration: January 2022 – Present

- Average of all modules till date is 78.6%

## SKILLS AND TOOLS:

---

**Programming Languages:** Python, R

**Databases:** MySQL, Oracle DB, Mongo DB

**Dashboarding Tools:** PowerBI

**Deployment:** Heroku Cloud Platform, Streamlit Cloud, Git version control

**Miscellaneous:** Pandas, Numpy, Matplotlib, Plotly, Tensorflow, Flask, Beautiful Soup, Selenium, Excel, Natural Language Processing, Computer Vision, Sklearn

## KEY PROJECTS IN DATA SCIENCE

---

### 1. Furniture Recommender WebApp: [\[GitHub Link\]](#)

Upload an image and this WebApp will recommend chairs, couches, beds and tables

- Implemented ResNet-50 to extract features from each image in the dataset.
- These features were fed to KNN model which by brute-forcing the Euclidean distance between the image uploaded by the user and all the images in the dataset, returned the indexes of 5 nearest neighbors of the image inputted by the user on the WebApp.
- At last, the images in the dataset corresponding to these 5 indexes were shown as recommendations on the WebApp.

### 2. Sales Insights Dashboard (PowerBI): [\[GitHub Link\]](#)

- Connected and transformed the raw data
- Built a relational data model
- Cleaned the data and created new calculated columns using Power Query Editor
- Designed an interactive dashboard to analyze sales, revenue, profits, returns and visualized insights about the product level trends, market trends in different regions, sales trends in different domains and identified high value customers and products.

### 3. WhatsApp Chat Analyzer: [\[WebApp Link\]](#) [\[GitHub Link\]](#)

This WebApp utilizes a series of data cleaning and data pre-processing pipelines to analyze the WhatsApp messages and deliver insights about the trends in user activity and user behaviour using interactive plotly charts and perform sentiment analysis not only with respect to the WhatsApp group as a whole but also with respect to each person in that group. This WebApp supports both 12 hour and 24 hour clock formats and also dd/mm/yyyy, mm/dd/yyyy and yyyy/mm/dd date formats. This WebApp was deployed on Heroku cloud platform.

#### 4. Web Scraping Projects:

- Text Web Scraping WebApp: [\[WebApp Link\]](#) [\[GitHub Link\]](#)  
This project takes a product name from the user and then searches for various product listings on India's largest online shopping site: Flipkart.com. It then scrapes customer reviews from each product listing and displays the latest 10 reviews in a tabular format on a webpage deployed on Heroku Cloud Platform.
- Google Image Scraper: [\[GitHub Link\]](#)  
This project uses Selenium to directly scrape images from google and stores them in their respective folders based on the keywords searched.

#### 5. Book Recommender WebApp: [\[WebApp Link\]](#) [\[GitHub Link\]](#)

This is a content-based recommender system where user has to select a book title from a list of 4000 books, then WebApp lists top 5 similar books based upon the description of the selected book.

- Description of each book was converted to numerical vectors using techniques such as BOWV, TF-IDF and Word2Vec algorithm
- Word2Vec (neural network based) which uses semantic meaning of words proved out to be the best one for this use case
- After pre-processing, Euclidean distance was measured to find pairwise similarity.
- The top 5 books who's description had the closest Euclidean distance with the book selected by the user on the WebApp was shown as recommendations.
- This WebApp was deployed on Streamlit Cloud Platform.

#### 6. Forecasting WebApp: [\[WebApp Link\]](#) [\[GitHub Link\]](#)

This WebApp can forecast any kind of numerical time series data using Facebook's prophet library and can also perform hyperparameter tuning. A user can tweak the following:

- **Horizon:** the time in future to forecast. It is expressed in days.
- **Seasonality:** choose between Additive seasonality or Multiplicative seasonality.
- **Trend components:** declare which trends want to discover and propagate. Daily should be selected if loading a dataset with hourly data. Weekly: Prophet will search for trend during days of the week (Monday to Sunday). Monthly: Prophet will search for trend during days of the month (1th to 31th). Yearly: will evaluate trend within months of the year (January to December)
- **Growth model:** choose between linear growth or logistic growth.
- **Holidays:** add holidays to the model. Available countries at the moment: Italy, Spain, France, United States, India, Germany and Ukraine.
- **Hyperparameters:** Change the scale of the changepoints or holidays. It impacts the flexibility of the model.

## CERTIFICATIONS & SPECIALIZATIONS (MOOCs)

---

### Coursera:

- Mathematics for Machine Learning (3 courses) | Imperial College London
- Machine Learning | Stanford University
- Applied Data Science with Python (5 courses) | University of Michigan
- Data Science with R (10 courses) | John Hopkins University
- DeepLearning.AI Tensorflow Developer Certification
- Google Data Analytics Certificate

### Udemy:

- Advance Excel Certification