

# Ryan Heng

## Data Scientist

### Contact Details:

ryanheng99@gmail.com | +65 9696 9890 | Singapore  
[linkedin.com/in/ryan-heng-bc/](https://www.linkedin.com/in/ryan-heng-bc/) | [github.com/ryanheng99](https://github.com/ryanheng99)

### Professional Summary

Data analyst with expertise in statistical modeling and analytical problem-solving. Adept at leveraging data to diagnose business challenges, optimize processes, and uncover market inefficiencies, enabling organizations to make informed and high-impact decisions.

### Work Experience

- **Data Scientist Intern** - Sanofi Manufacturing, Singapore  
May 2025 - Apr 2026, (1 Year)
  - Built and maintained ETL data pipelines integrating MES, DCS, and historian data, enabling reliable access to high-frequency manufacturing datasets for analytics and modeling
  - Performed exploratory data analysis on process and equipment data to identify trends, anomalies, and key drivers impacting production performance
  - Developed and evaluated time-series models (ARIMA, LSTM) to forecast process variables and support early detection of equipment or process deviations
  - Implemented anomaly detection techniques to identify abnormal patterns in manufacturing data, improving monitoring of process stability and quality metrics
  - Collaborated with cross-functional teams (Engineering, Automation, Quality) to translate data insights into actionable improvements for yield, throughput, and downtime reduction
  - Developed data transformations and SQL queries in Snowflake and MSSQL to support scalable data storage and analytics workflows
  - Built dashboards and visualizations to communicate key metrics and insights to technical and non-technical stakeholders
- **Data Engineer Intern** - Power Hydraulics, Singapore  
Mar 2023 - Aug 2023 (6 Mths)
  - Calibrated and optimized PLC, SCADA, and DCS control logic systems for marine vessel ballast water systems ensuring structural integrity, and improve maneuverability
  - Built and improved batch streaming data pipelines for sensor and equipment telemetry
  - Automated test data analysis using Python and MATLAB scripts, reducing manual analysis effort and increasing test throughput for engineering validation workflows
- **Associate Researcher Intern** - Procter & Gamble, Singapore (SkinCare department)  
Aug 2019 - Mar 2020 (6 Mths)
  - Built and analyzed experimental datasets from skincare formulation trials using Python and Excel, applying statistical methods to identify key drivers of product stability and performance
  - Developed data-driven insights from laboratory and consumer testing data, enabling optimization of formulations based on efficacy, texture, and user experience metrics
  - Automated data cleaning and preprocessing workflows for experimental results, improving analysis efficiency and reducing manual errors in reporting

### Skills

- **Programming:** Python, SQL, MATLAB, VBA
- **Data Science & Analytics:** Machine Learning/Deep Learning(PyTorch, TensorFlow, Scikit-Learn) Time-Series Analysis, Anomaly Detection, Statistical Analysis, Data Visualization(Matplotlib, Plotly)
- **Data Engineering:** ETL Pipelines, Data Warehousing(Snowflake), Data Integration
- **Version Control:** Git, GitHub, CI/CD
- **Library & Tools:** Pandas, NumPy, Matplotlib, Plotly, PyTorch (basic), TensorFlow (basic)

### Awards & Certifications

- Sanofi:
  - Certification in AWS Data Analytics (Nov 2025)
  - Certification in Datalake IA - Cloud Data Engineer Project & Operations (Nov 2025)
  - Certification in Advanced knowledge of AWS Console & AWS Services (Oct 2025)
- Wilmar International Scholar 2023

### Projects

- **Smart MFG OEE Data Pipeline for analysis** (Sanofi Smart Manufacturing)
  - Developed data pipelines to process manufacturing data and compute OEE metrics (availability, performance, quality) for production analysis
  - Implemented data transformations and KPI calculations to support downtime and efficiency tracking
  - Produced structured, visualization-ready datasets for dashboarding and reporting
  - **Tools: Python, SQL, RestAPI, Plotly, Matplotlib**
- **Time-Series Forecasting & Anomaly Detection** (Sanofi Smart Manufacturing)
  - Built and evaluated time-series models (ARIMA, LSTM) to forecast high-frequency manufacturing data and support early identification of process deviations
  - Implemented statistical anomaly detection methods (e.g., Z-score) to identify outliers in process and equipment data
  - Performed data preprocessing and feature engineering on time-series datasets
  - **Tools: Python, Pandas, NumPy, Statsmodels, PyTorch, Matplotlib**
- **Image Classifier & Regression ML Models**
  - Developed CNN-based models for image classification tasks and regression models for structured/tabular data
  - Conducted model evaluation, comparison, and data preprocessing to improve performance
  - **Tools: PyTorch, TensorFlow, Scikit-Learn, Matplotlib**

### Education

- **BEng Chemical Engineering (Specialised in Data Engineering) with Honours (2nd Upper)** - Singapore Institute of Technology-Technical University of Munich, Singapore (2022-2026)

## Languages

---

- English (Fluent)
- Mandarin (Fluent)