

Che-Wei (Jarvis) Lee

617-286-9011 | lee.chew@northeastern.edu | [linkedin.com/in/che-wei-lee-6b377a276](https://www.linkedin.com/in/che-wei-lee-6b377a276) | Boston, MA

SUMMARY

MS in Data Analytics Engineering candidate who builds self-updating analytics products (two currently live) that turn messy data into insight with SQL, Python, and Tableau. Seeking a Fall 2026 data analytics co-op (F-1 CPT, no sponsorship required).

EDUCATION

Northeastern University (Boston, MA)	Expected May 2027
Master of Science in Data Analytics Engineering	
Relevant Coursework: Machine Learning, Computation and Visualization, Data Management	
National Changhua University of Education (Changhua City, Taiwan)	Jan 2025
Bachelor of Business Administration	

TECHNICAL SKILLS

Programming & Query: Python, SQL
Data Analytics: Data Cleaning, EDA, Statistical Analysis, A/B Testing & Experimentation, Time-Series & Geospatial Analysis
Visualization & BI: Tableau, Power BI, Excel, Plotly, Matplotlib
Databases & Cloud: MySQL, SQLite, Google BigQuery; AWS/Azure/GCP (market & pricing analytics)
Tools: Git, GitHub Actions (CI/CD), Jupyter, AI-assisted development (Claude Code)

EXPERIENCE

Chunghwa Telecom (Taipei, Taiwan)	Jan 2024 - Jun 2024
<i>Business Analyst Intern, Cloud Team</i>	
- Analyzed monthly usage and traffic data across 3,000+ AWS customer accounts in Excel and Power BI, delivering monthly usage reports to the marketing and product managers to guide account priorities. - Evaluated YouTube vs. Google Ads for a Microsoft 365 campaign with scraped audience and pricing data, informing the final two-channel strategy; monitored weekly ad-performance reports. - Presented 12+ biweekly competitive-intelligence reports on AWS, Azure, and GCP to my mentor and the department head; benchmarked ChatGPT, Claude, and Copilot via cross-scoring as LLM tools emerged in early 2024.	

PROJECTS

Multi-Cloud AI Infrastructure Market Analysis Python, SQL, BeautifulSoup, Tableau, GitHub Actions	2026
- Collected and cleaned 2M+ raw AWS, Azure, and GCP pricing records into a curated ~12K-record dataset (134 regions, 9 service categories, 17+ accelerator types), modeled as a SQLite star schema. - Wrote 8 SQL analyses using window functions and self-joins (week-over-week price moves, YoY growth via LAG, spot-discount matching, regional price dispersion) to track GPU pricing across ~200 cloud regions. - Automated a weekly GitHub Actions pipeline that appends ~5,000 GPU price points per run to a live public dashboard, and delivered 3 Tableau dashboards — surfacing GCP's market-share growth from 11% to 15% over 6 quarters.	
U.S. Aerospace Innovation Atlas Python, GeoPandas, Plotly Dash, BigQuery (SQL)	2025
- Built an end-to-end pipeline mapping 16 years (2000-2015) of U.S. aerospace patents (5,200+) across 300+ metro areas, integrating USPTO, U.S. Census, and Google BigQuery sources. - Diagnosed and fixed a silent data-quality bug that had zeroed out Los Angeles — the #1 metro with 62 patents — by replacing name-matching with stable CBSA-code spatial joins; validated against official USPTO totals. - Engineered concentration (Herfindahl index, top-5 share), shift-share growth, and per-100k metrics; shipped an interactive Plotly Dash dashboard and extended to 38K+ patents via BigQuery SQL (aviation vs. space).	
NFL Offensive Performance Analytics Python (pandas), Power BI (DAX), Plotly	2025
- Aggregated ~22.6K nflverse player-week records (2021–2024, 600+ players, 32 teams) into 2,345 season-player rows. - Engineered 4+ efficiency metrics (EPA, target share, catch %, completion %) across QB/RB/WR/TE to measure per-play value — revealing top-yardage players are often not the most efficient. - Built 2 dashboard versions — Power BI (explicit DAX measures, KPI cards, season/position slicers) and interactive Plotly — pairing every volume-leader view with a matching efficiency view.	