WORK PORTFOLIO

ANASTASIIA GOLOVCHENKO

Financial Analyst | MS Data Science | PhD Economics

ANASTASIIA@SASSA.COM

+13054988651

LinkedIn

OVERVIEW

TECHNICAL SKILLS & EXPERTISE

Domain Expertise:

- Financial Markets
- Risk Assessment
- Market Analysis
- Economic Research
- Quantitative Finance
- Econometrics
- Health Economics

Analytics & Visualization:

- Statistical Analysis
- Financial Modeling
- Panel Data Analysis
- Machine Learning
- Big Data Analytics
- Tableau, Power BI
- Geospatial Analysis

Technical Tools:

- SQL
- Python, R
- Stata
- Google Cloud Platform
- Apache Spark, PySpark
- Data Engineering

Quantitative Finance

FINANCIAL MODELING & VALUATION ANALYSIS: TESLA INC

Equity valuation of Tesla Inc using multiple valuation methodologies including DCF analysis, Enterprise Value multiples, and Trading Comparables analysis.

Methodology:

Step 1: Revenue Forecasting

- Evaluated 3 forecasting approaches: historic growth rate, market growth rate, research forecasts
- Selected research forecasts method (16.9% average growth) based on professional analyst consensus from major investment banks

→

Step 2: Financial Projections & DCF Model

- Built 5-year financial projections with income statement modeling
- Calculated WACC (weighted average cost of capital): 8.31%
- Developed discounted cash flow (DCF) model
- Conducted sensitivity analysis across WACC and perpetuity growth rate assumptions

Excel Modeling

DCF Analysis

Step 3: Enterprise Value Analysis (EV)

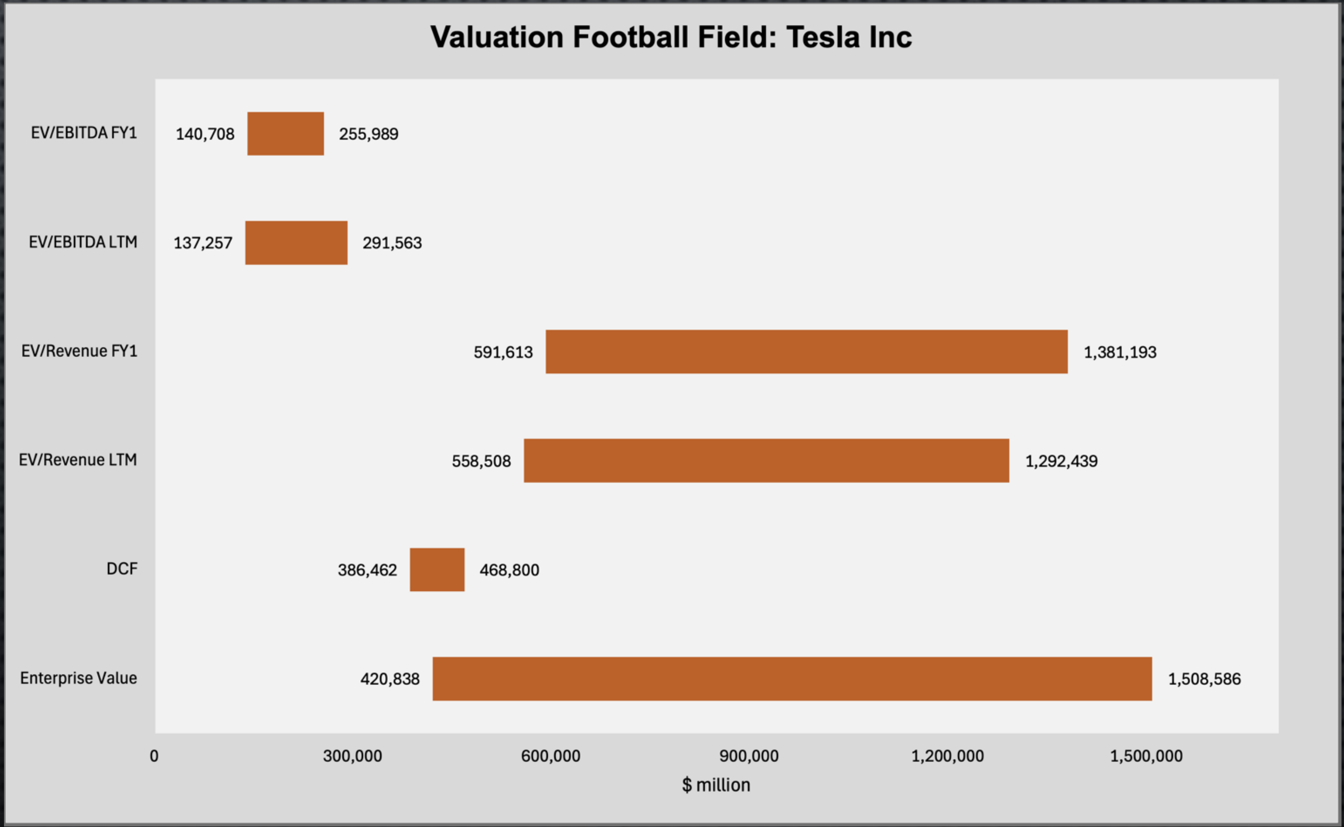
- Calculated current Enterprise Value based on share price at close of March 16th, 2025
- Computed Tesla's valuation multiples vs LTM and FY1 metrics
- Established 52-week Enterprise Value trading range for context

Step 4: Trading Comparables Analysis

- Benchmarked vs automotive/EV peers: BYD, Rivian, GM, Volkswagen, etc.
- Applied peer EV/Revenue and EV/EBITDA multiples to Tesla metrics

Step 5: Valuation Summary (Football Field)

- Generated valuation range across multiple approaches



View Full Financial Model

Quantitative Finance

FAMA-FRENCH 3-FACTOR MODEL IMPLEMENTATION

Advanced asset pricing and risk factor analysis using Python for quantitative portfolio attribution and performance measurement across equity securities.

Methodology:

Step 1: Data Collection & CAPM Baseline

- Multi-source data integration: Yahoo Finance equity prices, Fama-French research database for risk factors
- Time series alignment: Resolved mismatched frequencies between daily prices and monthly factor data using pandas resampling
- CAPM regression: Microsoft vs S&P 500 analysis yielding $R^2 = 0.534$ and Beta = 0.90

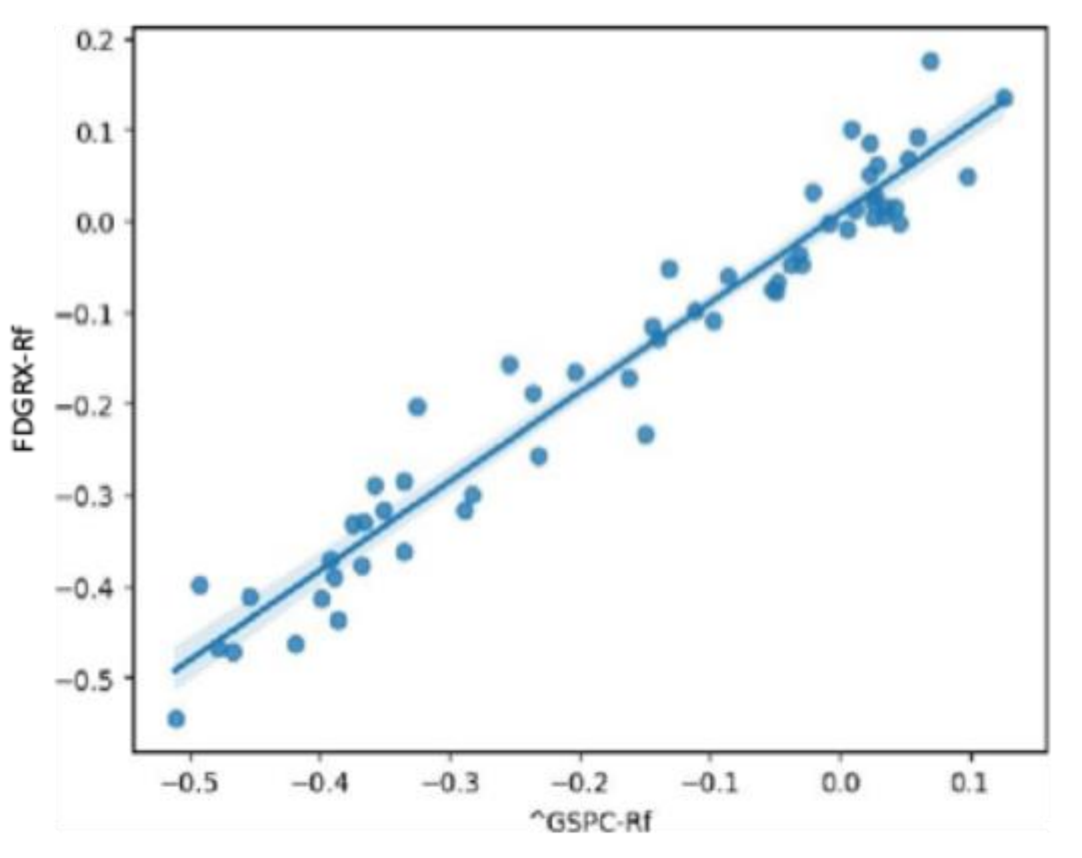
Step 2: 3-Factor Model Enhancement

- Excess return calculation: Risk-adjusted returns using Fama-French risk-free rate
- Dramatic model improvement: R^2 increased from 0.534 to 0.948
- Beta refinement: Microsoft beta = 0.98 (near market sensitivity)
- Statistical validation: F-statistic improved from 66.5 to 1030

Statistical Modeling | Risk Assessment | Python

STEP 3: FIDELITY FUND APPLICATION & INVESTMENT INSIGHTS

- Applied three-factor model to Fidelity Growth Company Fund (FDGRX) adding Size (SMB) and Value (HML) factors
- Exceptional model fit: $R^2 = 0.951$ (95.1% explanatory power)
- Positive alpha of 0.64% monthly indicating risk-adjusted outperformance
- Market beta = 1.15 (fund is more volatile than market)
- SMB = 0.15 (slight small-cap tilt - benefits when small caps outperform)
- HML = -0.44 (strong growth preference over value stocks)



FDGRX EXCESS RETURNS REGRESSION

View Full Model

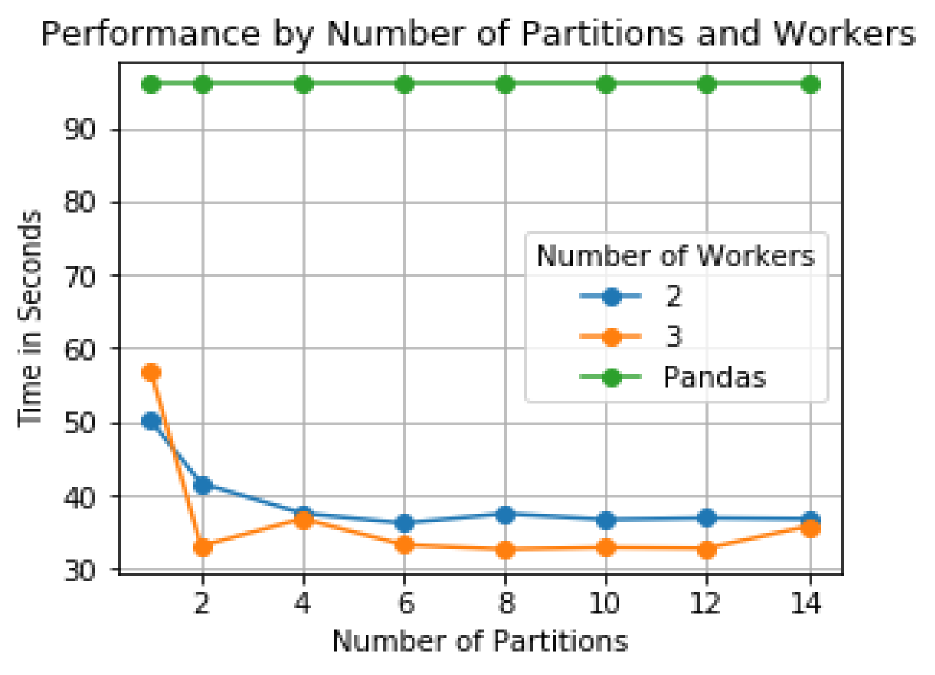
DISTRIBUTED SENTIMENT ANALYSIS: GEOPOLITICAL NEWS COVERAGE

Large-scale sentiment analysis of YouTube video transcripts, focusing on geopolitical content analysis across multiple news channels (BBC, CNN, SABC News, Al Jazeera English).

Methodology:

STEP 1: DATA EXTRACTION & PIPELINE DESIGN

- Extracted video URLs from 4 news channels via Google YouTube Data API v3
- Built distributed data pipeline using Apache Spark RDDs for parallel processing of audio extraction, standardization, and cloud storage upload
- Performance optimization: reduced processing time by 66% using Spark vs. Pandas (96s → 32.56s for 15 videos)

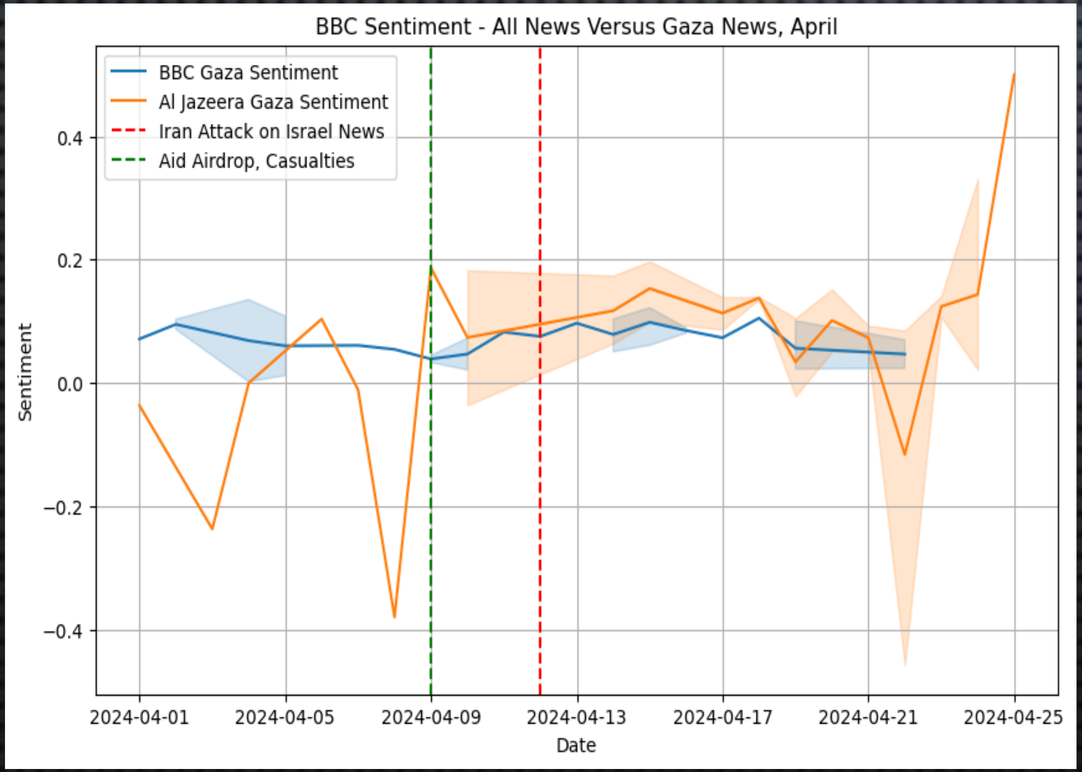


PERFORMANCE OPTIMIZATION: SPARK VS PANDAS

Apache Spark & PySpark | Google Cloud Platform | NLP | Data Engineering

Step 2: Speech-to-Text Transcription

- INTEGRATED GOOGLE CLOUD SPEECH-TO-TEXT API FOR AUDIO TRANSCRIPTION
- ACHIEVED 88.74% ACCURACY FOR SINGLE-REPORTER NEWS VIDEOS (11.26% WER)



SENTIMENT VARIATIONS: CROSS-CHANNEL ANALYSIS

Step 3: Sentiment Analysis & Insights

- APPLIED TEXTBLOB MODEL TO ANALYZE SENTIMENT SCORES OF GEOPOLITICAL EVENTS
- CORRELATED SENTIMENT SHIFTS WITH REAL-TIME EVENTS: GAZA AID INCIDENT, IRANIAN ATTACKS ON ISRAEL
- IDENTIFIED EDITORIAL VARIATIONS BETWEEN BBC AND AL JAZEERA: DIVERGENT COVERAGE OF ISRAEL-IRAN CONFLICT

[View Full Project]

Big Data Analytics & Machine Learning

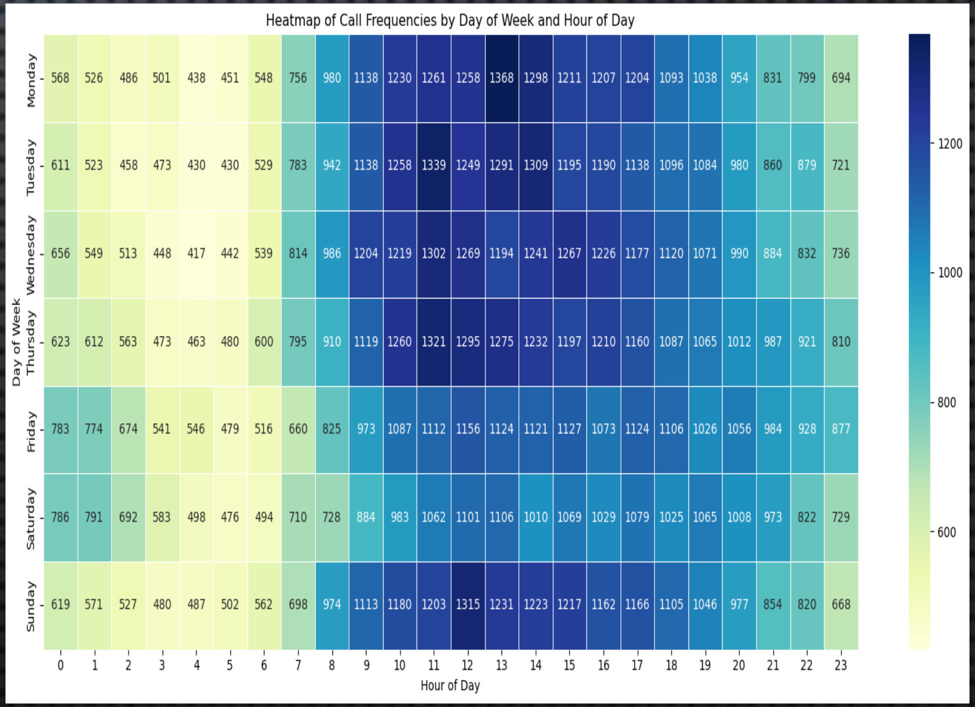
EMERGENCY RESPONSE TIME PREDICTION: SAN FRANCISCO FIRE DEPARTMENT

Predictive modeling of emergency response times using machine learning and geospatial analysis across 152,384 emergency incidents to optimize service delivery.

Methodology:

STEP 1: DATA ENGINEERING & GEOSPATIAL INTEGRATION

- Processed 2.6GB dataset (20+ years) using Stata, focusing on 2022 to create 152,384 unique emergency records
- Integrated OSRM routing API and web-scraped fire station locations for driving distance calculations across 44 fire stations
- Engineered temporal features and response time target using San Francisco County's official definition
- Validated geospatial accuracy revealing strong correlation between emergency call frequency and population density patterns



TEMPORAL PATTERN ANALYSIS: EMERGENCY CALL FREQUENCIES

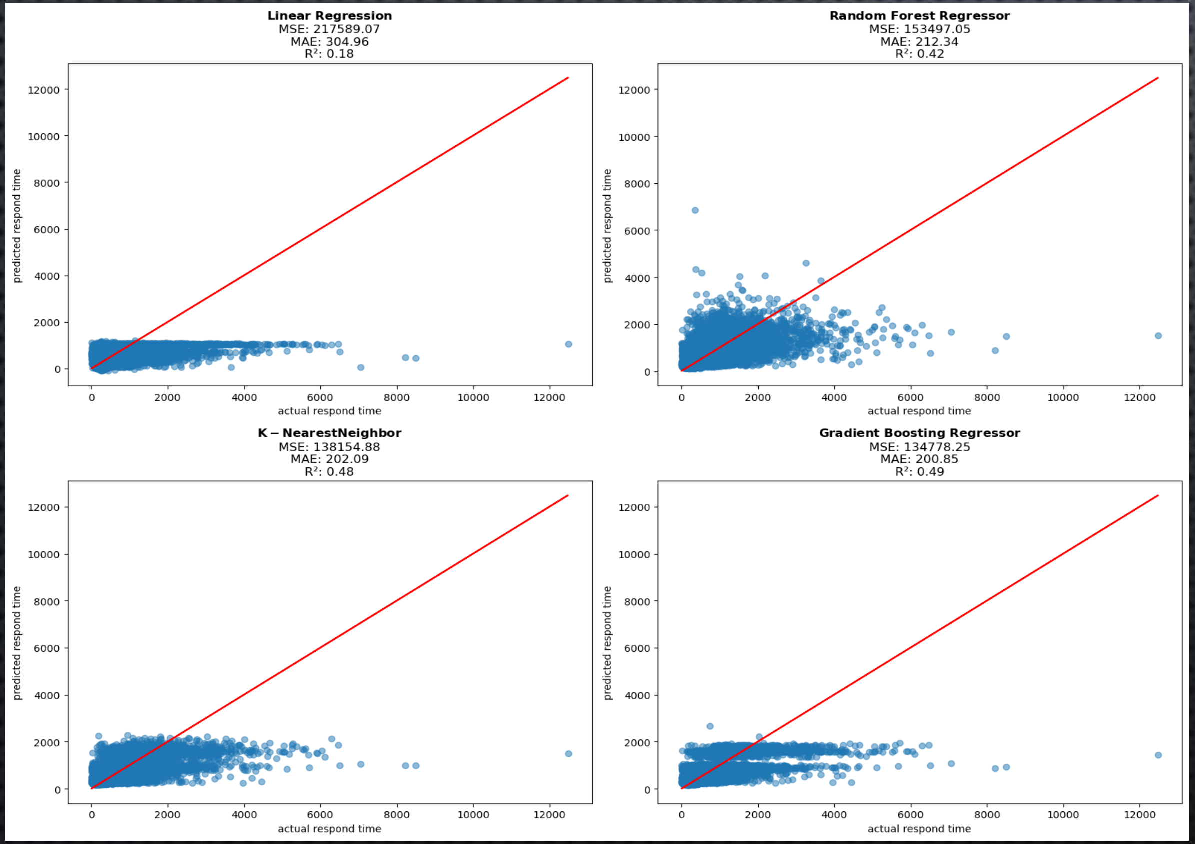
Python

Data Engineering

Geospatial Analysis

STEP 2: PREDICTIVE MODELING

- Trained regression (Linear Regression, Random Forest, Gradient Boosting, KNN) and classification models (Random Forest, SVM, KNN, AdaBoost) across 11 features
- Best performance: Gradient Boosting ($R^2 = 0.49$, MAE = 201 seconds); classification accuracy 24-28%
- Key finding: driving distance showed minimal correlation (0.01) with response times - operational factors dominate



MODEL COMPARISON: ACTUAL VS PREDICTED RESPONSE TIMES

STEP 3: INSIGHTS & IMPLICATIONS

- Emergency calls peaked during 10-17h daily and weekends with response times increasing 35% during peak periods
- Geographic hotspots identified: highest call volumes in Tenderloin, Mission, and Bayview neighborhoods
- Life-threatening calls met San Francisco's 10-minute policy target approximately 90% of the time
- Response time prediction proves challenging due to unmeasured operational factors; interval-based estimates recommended

View Full Project

Risk Assessment

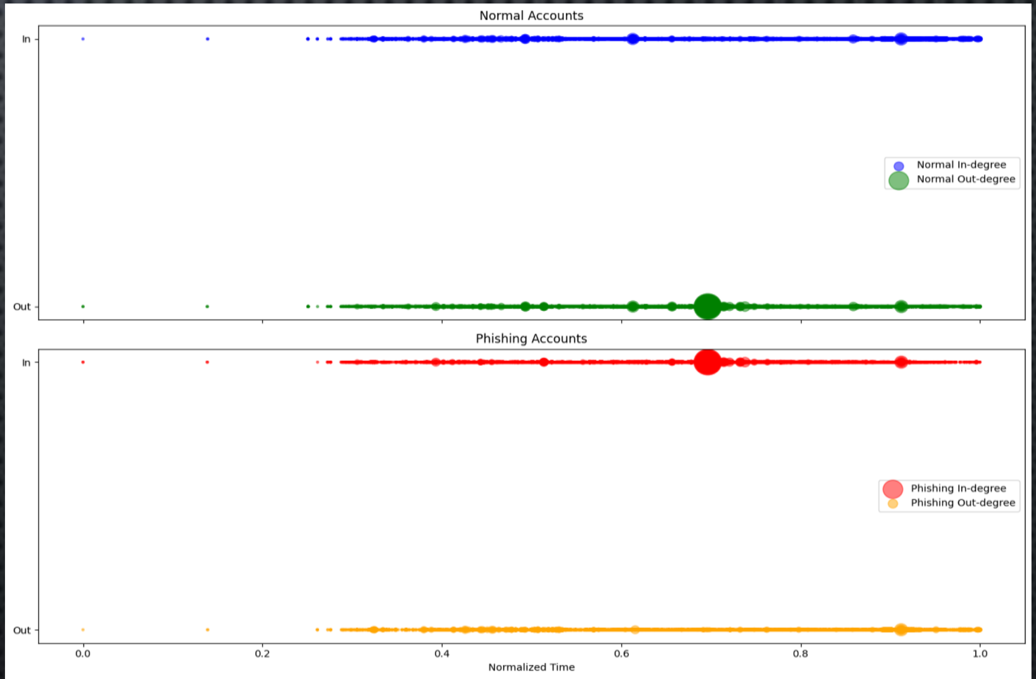
ETHEREUM BLOCKCHAIN FRAUD DETECTION

Applied graph representation learning to detect anomalous wallets in Ethereum blockchain transactions using node embeddings and GCNs.

Methodology:

STEP 1: DATA ENGINEERING & GRAPH CONSTRUCTION

- Constructed Ethereum transaction graph from 1,600+ verified phishing addresses with 30,000+ nodes (wallet addresses) and 80,000+ edges (ETH transfers) weighted by transaction value
- Engineered features: structural (node degree, clustering coefficient), transactional values, and interaction intensity metrics
- Validated distinct behavioral patterns between legitimate and fraudulent accounts through temporal analysis



TRANSACTION BEHAVIOR PATTERNS: NORMAL vs PHISHING

Machine Learning

Network Analysis

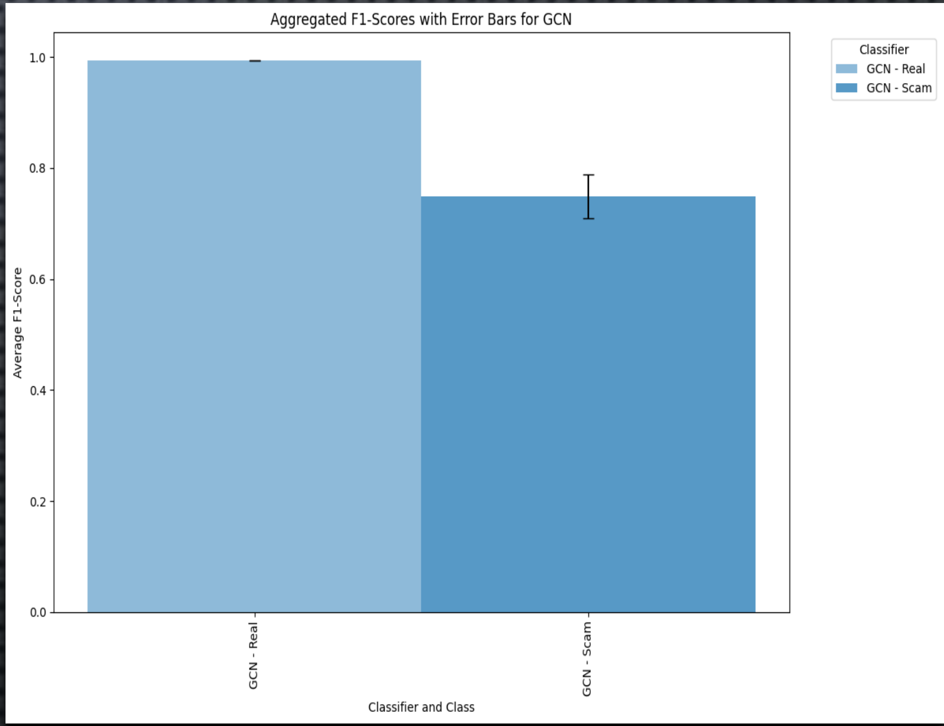
Python

STEP 2: EMBEDDING & CLASSIFICATION

- Implemented full-graph methods: Node2Vec, Graph2Vec, GraphSAGE, and Graph Convolutional Networks (GCN)
- Applied ego-graph approaches using k-hop subgraph methods for local neighborhood fraud detection
- Used SMOTE techniques to address extreme fraud class imbalance
- Trained classifiers: Logistic Regression, Random Forest, SVM, KNN across 10 experimental iterations

STEP 3: APPLICATION & INSIGHTS

- GCN achieved best performance with 98.87% accuracy, 99.53% precision (legitimate), 79.37% recall (fraud), outperforming traditional embedding methods (Node2Vec: 95.5%, GraphSAGE: 98.0%)
- Ego-graph approaches effectively captured localized fraud patterns with reduced computational requirements
- Developed scalable fraud detection pipeline for cryptocurrency exchanges



GCN F1-score by class (Real vs Scam)

View Full Project

Econometrics

PREDICTIVE MODELING OF SANCTIONS ON RUSSIAN OLIGARCHS

Econometric analysis of sanctions patterns on 200 wealthiest Russians following the 2022 Ukraine invasion, identifying key determinants through predictive modeling.

Methodology:

STEP 1: DATABASE CONSTRUCTION & VARIABLE ENGINEERING

- Cross-referenced Forbes "200 Richest Russians" with OpenSanctions database
- Engineered 7 binary predictive variables: Kremlin proximity, strategic industry involvement, emigration status, wealth ranking (top 50), Western investments, war opposition, and media exposure
- Manually verified political connections and business affiliations to ensure data accuracy

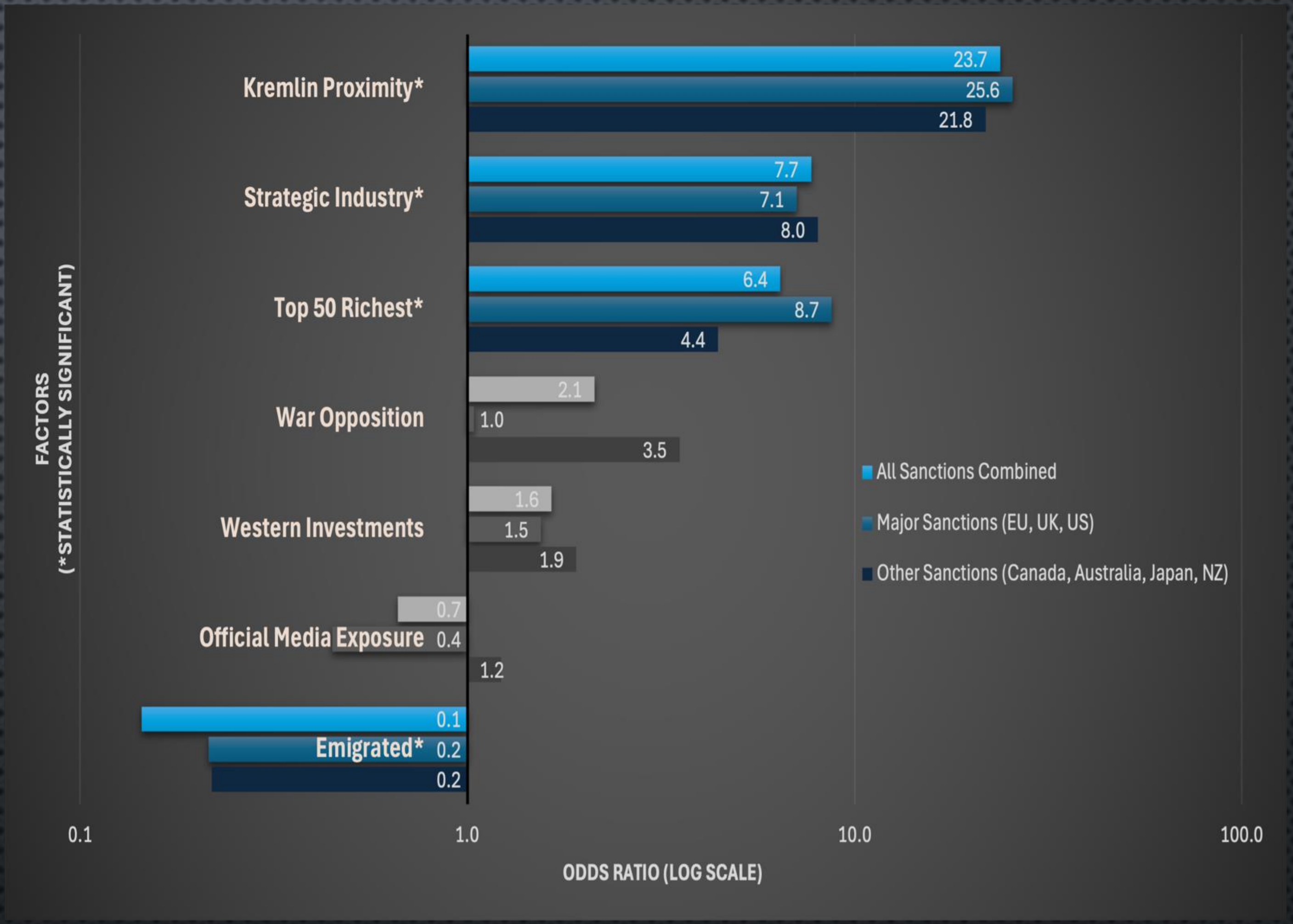
STEP 2: LOGISTIC REGRESSION MODELING

- Developed 3 logistic regression models: major sanctions (US/EU/UK), other sanctions (Canada/Australia/Japan/NZ), and all sanctions combined
- Achieved moderate to strong model fit with Pseudo R^2 of 0.43-0.49

Logistic Regression | Stata

Step 3: Key Findings & Policy Insights

- Sanctions concentrated among wealthiest: 68% of top 50 sanctioned vs 35% overall
- Kremlin proximity strongest predictor increasing odds of sanctions by 23.7 times, followed by strategic industry involvement (7.7x) and top 50 wealth status (6.4x)
- Emigration provided significant protective effect, reducing odds of sanctions by 84%
- War opposition, Western investments, and media exposure showed no statistically significant impact
- Research contributed to 3 additional publications on war profiteering, expropriation patterns, and Ukrainian reforms



Determinants of individual sanctions on Russian oligarchs

View Publication

View Stata Output

Health Economics

COVID-19 HEALTH SYSTEMS PERFORMANCE ASSESSMENT

Panel data analysis evaluating health system performance across 5 European countries (UK, France, Italy, Germany, Spain) in addressing the health and economic consequences of COVID-19

Methodology:

Step 1: Panel Data Engineering & Index Design

- Selected and justified 9 explanatory indicators to assess health system performance
- Constructed a quarterly panel dataset (Q1 2020–Q4 2022) for five EU countries by integrating Our World in Data, OECD, and Oxford COVID-19 Government Response Tracker
- Developed a composite Health System Performance Index (HSPI) as the dependent variable, combining min-max normalized GDP growth and inverted excess mortality (P-scores) using equal weighting

→

Step 2: Econometric Modeling

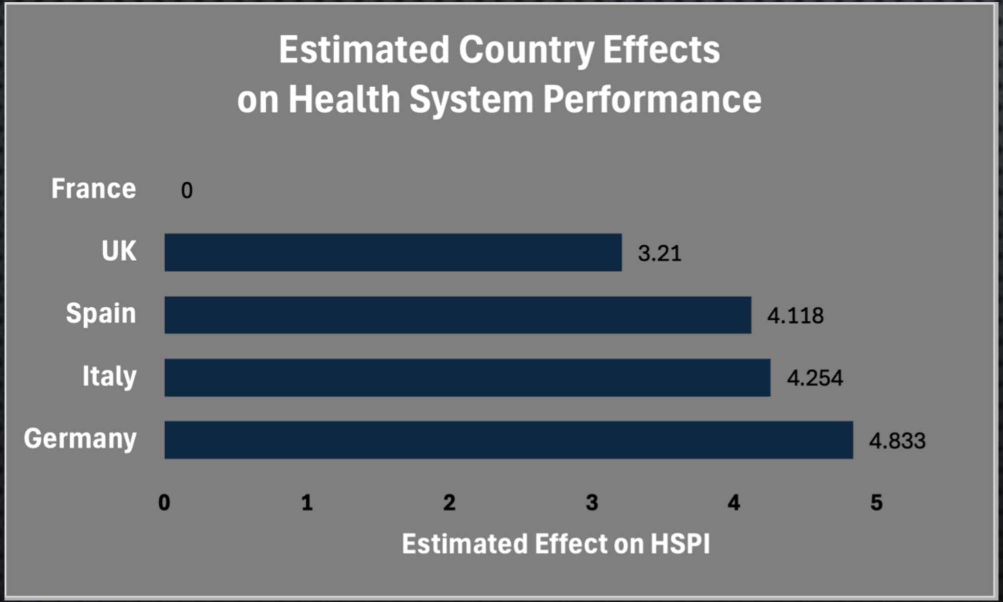
- Implemented three models: Pooled OLS, Panel Fixed Effects (FE) and Panel FE with Country Dummies
- Controlled for country-specific and time-specific effects; selected FE model over RE based on Hausman test results
- Achieved progressive model fit: R^2=0.292 (OLS) → 0.483 (FE) → 0.936 (Country Dummies)

Panel Data Analysis

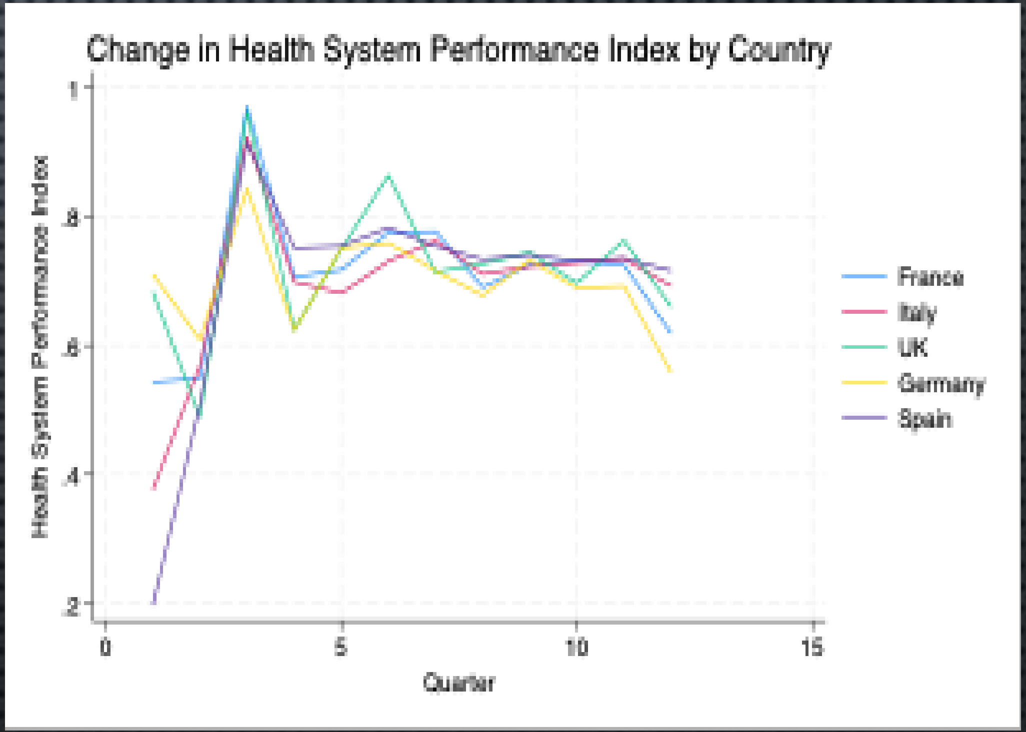
Stata

STEP 3: KEY FINDINGS & POLICY IMPLICATIONS

- Vaccination coverage, Testing capacity (tests per confirmed case) and consumer confidence Index had strong positive associations with health system performance.
- Unemployment, ICU occupancy, and case fatality rates negatively impacted performance
- Germany ranked highest in overall performance, followed by Italy, Spain, UK, with France trailed due to high ICU burden and fatality rates



Country Fixed Effects from Panel Regression Model



Health System Performance Index Evolution by Country.

View Full Project

Market Analysis

INTERNATIONAL DEVELOPMENT MARKET ANALYSIS

A market entry strategy for Spring Impact (UK-based NGO) expansion into East and Southern African markets

Methodology:

STEP 1: MARKET & COMPETITIVE ASSESSMENT

- Evaluated 23 countries using PESTLE framework to assess political stability, economic conditions, and infrastructure development
- Mapped 25+ competitors across local and global organizations, identifying market gaps in training programmes and partnerships
- Identified 16+ potential clients including NGOs, social enterprises, and development organizations

Economic Research

Risk Assessment

STEP 2: ENTRY STRATEGY & RECOMMENDATIONS

- Analyzed 5 entry models from franchising to wholly-owned operations
- Recommended dual-hub approach: Kenya for East Africa, South Africa for Southern region
- Developed implementation roadmap with franchising model and legal compliance framework

Key outcome:

Kenya and South Africa identified as optimal hubs with franchising model recommended for sustainable market entry and regional expansion.

COMPETITIVE LANDSCAPE ANALYSIS: SOUTHERN AFRICA

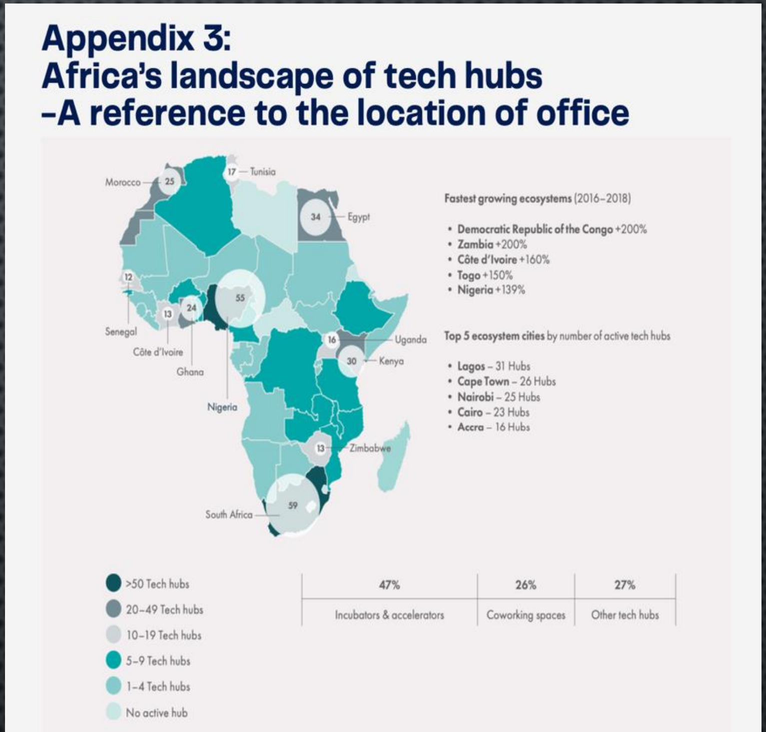
(iv) Analysis & Recommendation 13

PESTLE Analysis: Political, Economical, Social, Technological, Legal, Environmental

	Pros	Cons
Kenya	• **Political & Economical:** Relatively stable • **Economical:** rapidly growing with a sizable customer base • **Social:** high mobility – labour&client; higher reputation • **Technological:** Better infrastructure	• **Economical:** High competition – Challenges in attracting and retaining clients → Higher payroll & High operating costs and competitive pricing
Uganda	• **Political & Economical:** government encouraging startups → a growing customer base • **Economic:** high potential of becoming a large oil exporter • **Social:** English is the official language; Youthful population.	• **Technological:** Infrastructure challenges • **Political & Economical:** Political instability & scandals – uncertainties in government processes • **Social:** Shortage in skilled talents
Seychelles	• **Political & Economical:** Relatively stable • **Political:** easy to set up a branch • **Technological:** Better infrastructure	• **Economical:** highly reliant on tourism – less resilient • **Environmental:** limited market for consultancy + it is a group of islands locates relatively far from the EA coast – hard to access the network
Tanzania	• **Political & Economical:** a growing economy & customer base; ◦ Investment incentives. • **Political & Social:** Regulatory complexities can be both a challenge and an opportunity for Spring to develop speciality on.	• **Technological:** Infrastructure challenges • **Political:** Bureaucracy hurdles and regulatory complexities – uncertainties in government processes • **Social:** Shortage in skilled talents
Rwanda	• **Political:** Relatively stable • **Political & Economical:** Investment incentives • **Social & Technology:** Focus on innovation could open more opportunities	• **Technological:** Infrastructure challenges • **Economical & Social:** Limited domestic market + landlocked geography: less access to market and limiting growth opportunities • **Social:** Shortage in skilled talents

Based on the PESTLE Analysis, Kenya rated the highest and therefore is the most recommended location for Spring to set up a presence in the region.
The following section will elaborate on actionable insights on how to implement the model of franchising/affiliate with local partners in Kenya specifically.

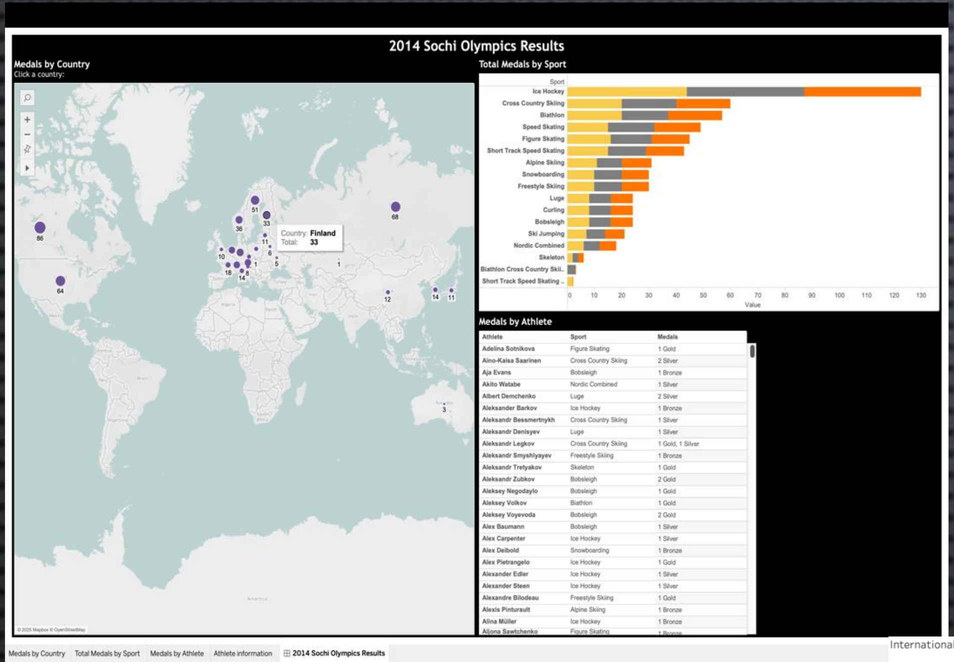
LOCATION SELECTION CRITERIA: EAST AFRICA



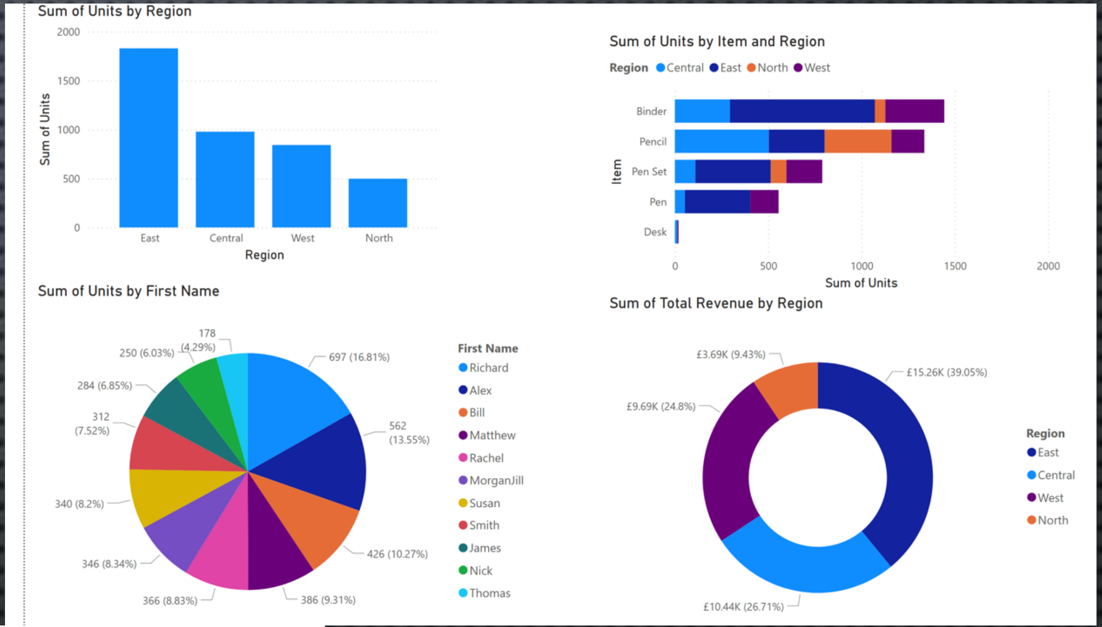
TECHNOLOGY ECOSYSTEM MAPPING: AFRICA

[View Full Project](#)

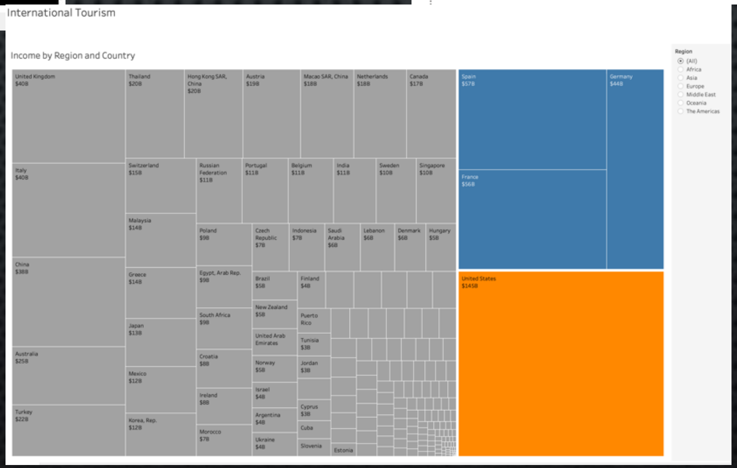
BUSINESS INTELLIGENCE & VISUALIZATION



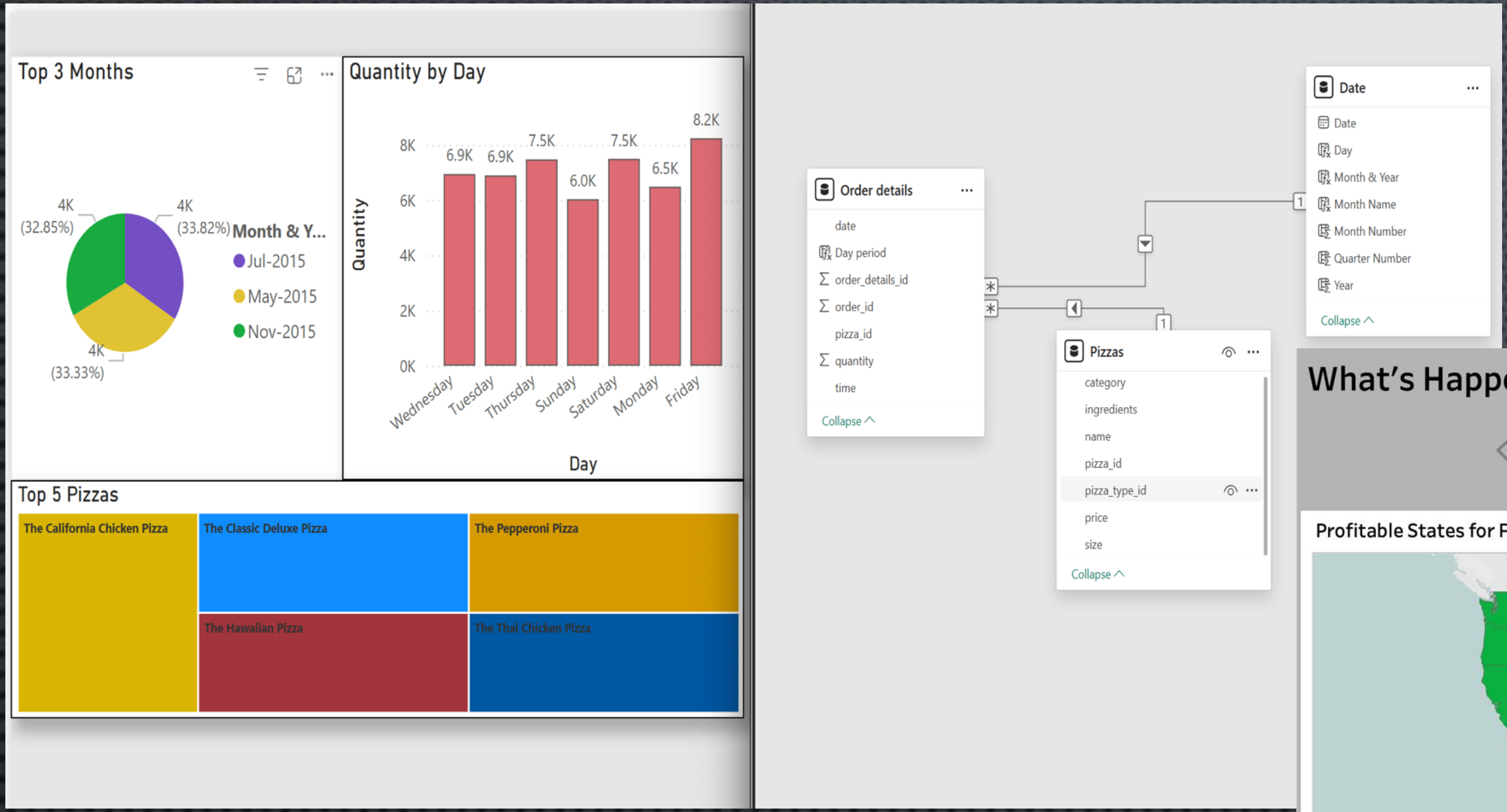
Olympic Performance Analytics (Tableau)



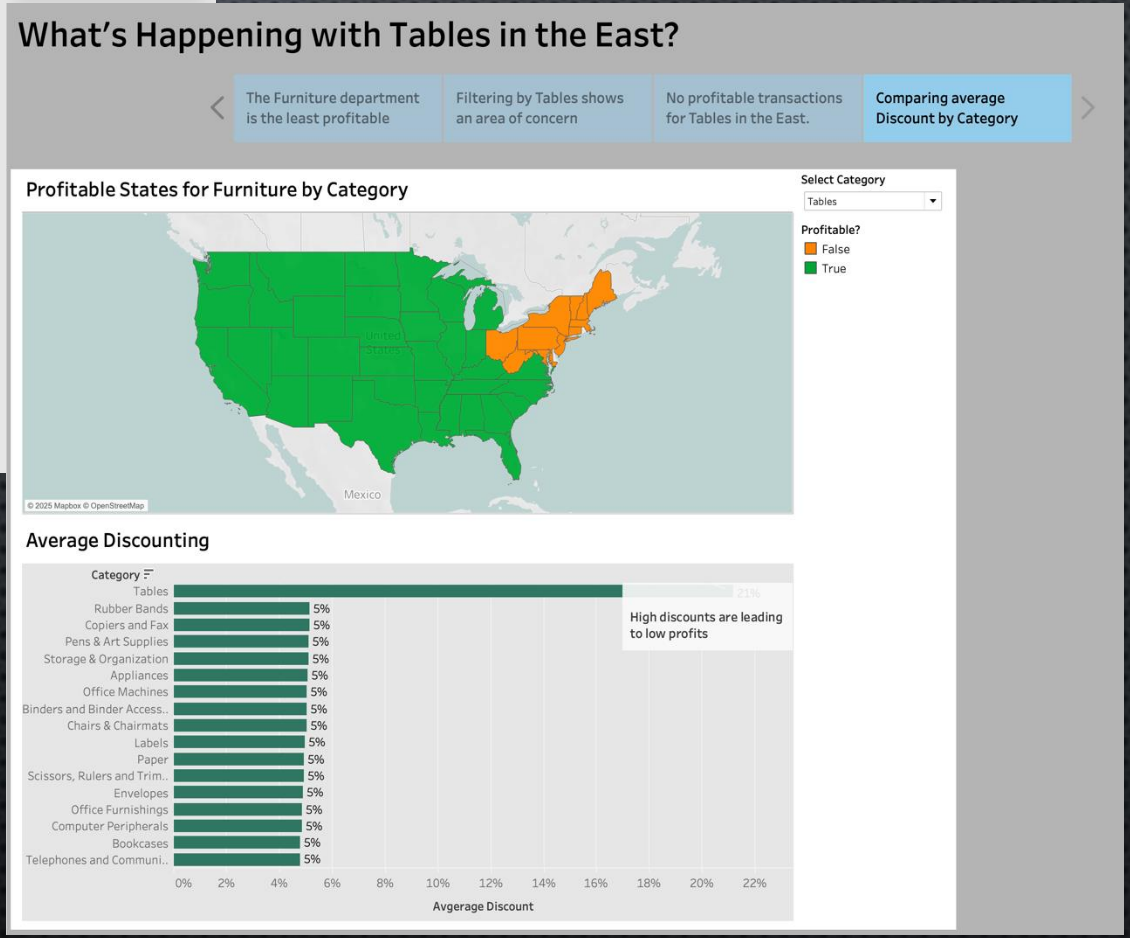
Supply Chain Analytics (Power BI)



Tourism Revenue Analysis (Tableau)



Sales Performance Dashboard (Power BI)



Business Intelligence Storytelling (Tableau)