**Course: Data Insights and Visualization (24 hrs / 8 sessions)**

**Integrated Tools:** Python (Pandas, scikit‑learn, statsmodels, Matplotlib), Power BI

**Datasets:**

* **Hands‑on Sessions:** NYC Property Rolling Sales
* **Final Project:** NYC Taxi Rides

**Core Learning Outcomes (LO):**

* **LO1:** Apply statistical and programming techniques to analyze complex datasets.
* **LO2:** Design and implement data visualizations using contemporary tools.
* **LO3:** Critically interpret data patterns and trends for decision‑making.
* **LO4:** Evaluate data quality, integrity, and ethics.
* **LO5:** Synthesise business intelligence for strategic problem‑solving.

**Session 1: Foundations & Environment Setup**

**Topics & Activities:**

1. **Importance of Data Analytics**: industry trends & case examples
2. **Types of Analytics**: descriptive, diagnostic, predictive, prescriptive
3. **Data Types & Formats**: numeric, categorical, datetime, spatial
4. **Data Quality & Key Characteristics**: accuracy, completeness, consistency
5. **Data Collection, Ethics & GDPR**: privacy, consent, anonymization
6. **Tool Setup**: Anaconda/Jupyter, Power BI installation
7. **Python Essentials**: lists, sets, dicts, tuples
8. **Hands‑on:** Pandas exploration on sample CSV

**Learning Outcomes:**

* LO1: Import and summaries data with Pandas.
* LO4: Identify and discuss ethical considerations in data collection.

**Session 2: Statistical Inference & Data Cleaning**

**Topics & Activities:**

1. **Descriptive Statistics**: central tendency, spread, skewness, CLT
2. **Sampling & Inference**: confidence intervals, standard error
3. **Hypothesis Testing**: t‑tests, p‑values, type I/II errors
4. **Correlation Concepts**: Pearson, Spearman
5. **ETL & Data Cleaning**: missing values, outlier detection
6. **Hands‑on:** Python exercises: compute stats, clean a messy dataset

**Learning Outcomes:**

* LO1: Implement inferential tests in Python.
* LO3: Interpret statistical results for hypothesis decisions.
* LO4: Execute ETL steps to produce clean, analysis‑ready data.

**Session 3: Exploratory Data Analysis on Property Sales**

**Topics & Activities:**

1. **Descriptive Stats**: aggregations in Pandas
2. **Univariate & Bivariate Analysis**: histograms, boxplots, scatter plots
3. **Correlation & Heatmaps**: compute and visualize relationships
4. **Time Series EDA**: resampling (monthly/quarterly), trend/seasonality plots
5. **Regression Overlays**: add best‑fit lines to scatter

**Hands‑on:**

* Notebook: EDA for assigned.

**Learning Outcomes:**

* LO1: Use Pandas & Matplotlib for comprehensive EDA.
* LO2: Build interactive EDA visuals in Power BI.
* LO3: Derive insights and articulate patterns from plots.

**Session 4: Regression Modeling & Business Interpretation**

**Topics & Activities:**

1. **Linear Regression Theory**: OLS fundamentals
2. **Implement in Excel vs Python**: statsmodels & scikit‑learn
3. **Model Evaluation**: R², RMSE, k‑fold cross‑validation
4. **Business Interpretation**: translating coefficients into recommendations

**Hands‑on:**

* Build models to predict sale price from sqft & year built.
* Compare performance and present coefficient insights.

**Learning Outcomes:**

* LO1: Develop and validate regression models.
* LO3: Critically assess model fit and assumptions.
* LO5: Formulate strategic recommendations based on model outcomes.

**Session 5: Time Series Forecasting Techniques**

**Topics & Activities:**

1. **Time Series Concepts**: stationarity, autocorrelation
2. **Decomposition**: additive vs multiplicative
3. **Stationarity Tests**: ADF, KPSS
4. **Forecasting Models**: ETS, ARIMA
5. **Power BI Forecasting**: built‑in line‑chart forecasts

**Hands‑on:**

* Decompose and test stationarity on borough series.
* Fit ARIMA, produce 3‑month forecast.

**Learning Outcomes:**

* LO1: Conduct time series diagnostics and modeling in Python.
* LO3: Interpret forecast accuracy and patterns.
* LO5: Recommend actions based on forecast scenarios.

**Session 6: Data Modeling & Advanced Power BI**

**Topics & Activities:**

1. **Data Warehouse Schemas**: star vs snowflake
2. **Power Query & M Basics**: transformations, parameterization
3. **Model Relationships**: cardinality, direction, role
4. **Intro to DAX**: measures vs columns, SUM, AVERAGE, COUNT
5. **Dashboard Principles**: chart selection, accessibility, layout
6. **Storytelling Techniques**: tailoring to C‑suite vs technical audiences
7. **Geospatial Mapping**: shape files, heatmaps in Power BI

**Hands‑on:**

* Build a star schema in Power BI data model.
* Create DAX measures and an audience‑focused dashboard page.

**Learning Outcomes:**

* LO2: Design robust data models and DAX measures.
* LO5: Craft dashboards tailored to stakeholder needs.
* LO3: Use geospatial visuals to highlight spatial insights.

**Session 7: Dashboard Construction & Customization**

**Topics & Activities:**

1. **Dashboard Layout & UX**: navigation, drill‑through, bookmarks
2. **Custom Visuals**: importing and configuring marketplace visuals
3. **Performance Optimization**: aggregations, incremental refresh
4. **Advanced Interactivity**: parameter tables, tooltips

**Hands‑on:**

* End‑to‑end build of a Power BI dashboard on NYC Property Sales.
* Include bookmark navigation and drill‑through to detail pages.

**Learning Outcomes:**

* LO2: Implement advanced Power BI features for interactivity.
* LO3: Ensure performance and usability in dashboards.
* LO5: Present cohesive BI story through dashboard flow.

**Session 8: Capstone Kick‑off & End‑to‑End Workflow**

**Topics & Activities:**

1. **Full Analytics Workflow**: data ingest → analysis → visualization → storytelling
2. **Final Project Briefing**: NYC Taxi Rides assignment
3. **Defining Research Questions & KPIs**
4. **Deliverables & Grading Criteria**
5. **Group Formation & Timeline Management**

**Hands‑on:**

* In‑class draft of project plan: objectives, methods, expected outputs
* Peer feedback on draft proposals

**Learning Outcomes:**

* LO1: Plan and structure an end‑to‑end data analysis project.
* LO2: Integrate Python and Power BI deliverables cohesively.
* LO5: Define measurable KPIs and articulate evaluation criteria.