**1. Fundamentals of Data Science**

* **Definition**: The process of extracting knowledge and insights from data using statistical, mathematical, and computational techniques.
* **Workflow**:
  1. Problem definition
  2. Data collection
  3. Data cleaning
  4. Exploratory data analysis (EDA)
  5. Feature engineering
  6. Modeling
  7. Evaluation
  8. Deployment & monitoring

**📊 2. Statistics & Probability**

* **Descriptive Stats**: Mean, median, variance, standard deviation, etc.
* **Inferential Stats**: Hypothesis testing, confidence intervals, p-values.
* **Probability**: Bayes' theorem, probability distributions (normal, binomial, etc.)

**🧮 3. Mathematics**

* **Linear Algebra**: Vectors, matrices, eigenvalues — essential for ML algorithms.
* **Calculus**: Gradients, partial derivatives — used in optimization (e.g., gradient descent).
* **Optimization**: Loss functions, convex optimization.

**💻 4. Programming**

* **Languages**:
  + **Python** (most popular): Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch
  + **R**: Especially in academia and statistics-heavy roles
* **Concepts**: Data structures, loops, functions, APIs

**📚 5. Data Handling**

* **Data Wrangling**: Cleaning, transforming, dealing with missing values
* **Libraries**: Pandas, NumPy
* **Databases**: SQL (joins, group by, subqueries), NoSQL (MongoDB)
* **File Types**: CSV, JSON, Excel, Parquet

**📈 6. Data Visualization**

* **Tools**: Matplotlib, Seaborn, Plotly, Tableau, Power BI
* **Charts**: Histograms, scatter plots, box plots, heatmaps

**🤖 7. Machine Learning Basics**

* **Types**:
  + **Supervised**: Regression, classification (e.g., linear regression, decision trees)
  + **Unsupervised**: Clustering, dimensionality reduction (e.g., K-means, PCA)
  + **Reinforcement**: Learning via feedback (used in games, robotics)
* **Model Evaluation**: Accuracy, precision, recall, F1 score, confusion matrix, ROC curve

**🛠️ 8. Tools & Environments**

* **Jupyter Notebooks**, VS Code
* **Git/GitHub**: Version control
* **Virtual Environments**: pip, conda
* **Cloud Platforms**: AWS, GCP, Azure

**🚀 9. Model Deployment**

* **APIs**: Flask, FastAPI
* **Containers**: Docker
* **CI/CD**: GitHub Actions, Jenkins
* **Monitoring**: Logging, model drift

**🧠 10. Soft Skills & Mindset**

* **Communication**: Presenting findings clearly to non-technical audiences
* **Critical Thinking**: Question assumptions and validate models
* **Curiosity & Learning**: Staying updated with trends (e.g., generative AI, AutoML)