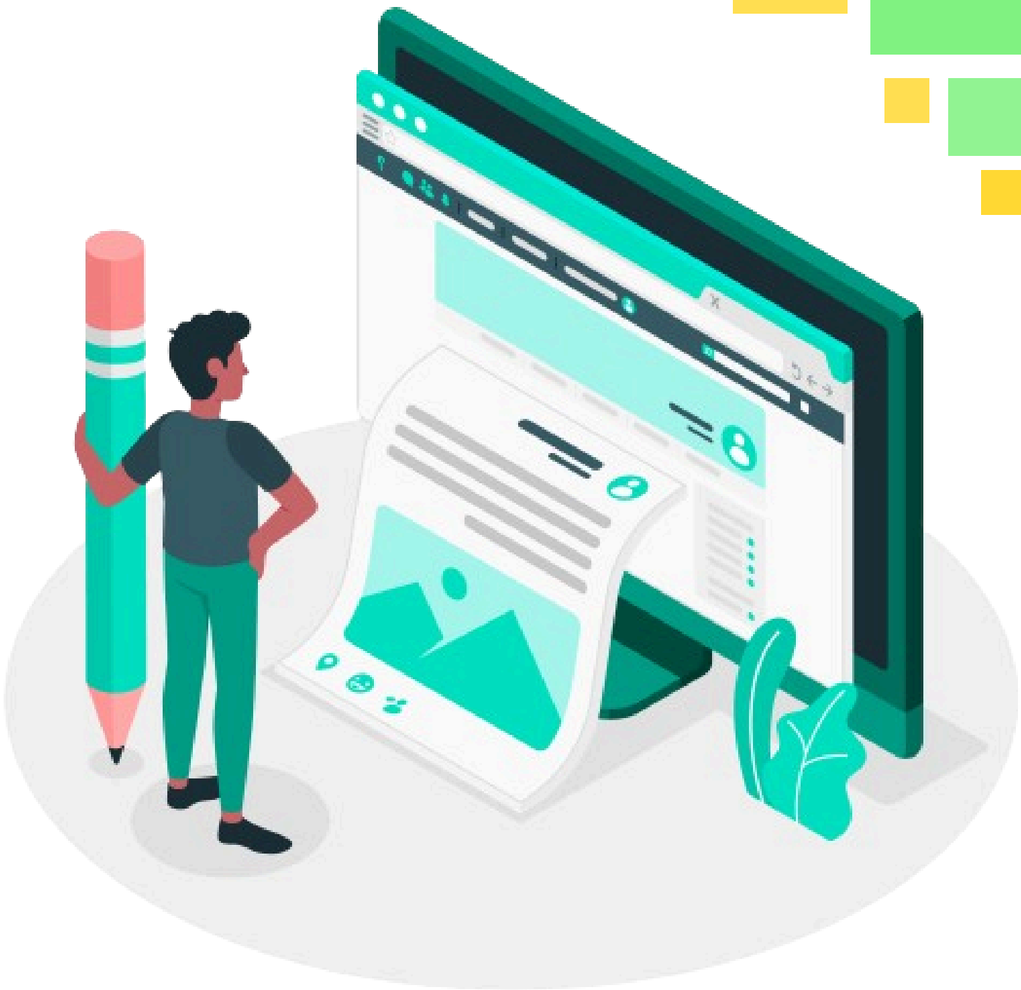
Codveda

Data Analytics

Content

1. About Us
2. Instructions
3. Submission
4. Task List



About Us

Welcome to Codveda Technology, where innovation meets excellence. Founded with a vision to empower businesses through cutting-edge IT solutions, we specialize in delivering tailored services that drive success in the digital era.

At Codveda, we offer a diverse range of services, including web development, app development, digital marketing, SEO optimization, AI/ML automation, and data analysis.

Our team of skilled professionals is committed to helping businesses unlock their full potential by providing innovative, scalable, and reliable solutions.

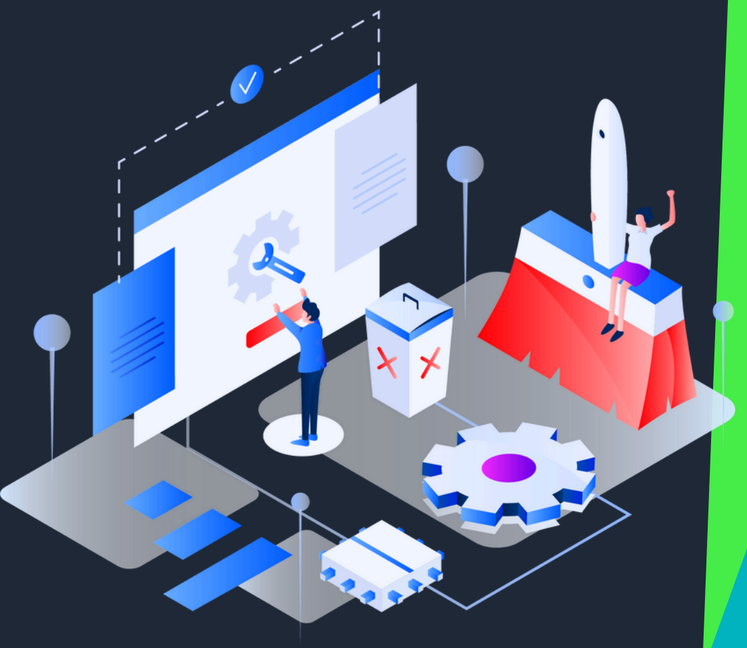
INSTRUCTIONS

- Update your LinkedIn profile with your achievements, including the offer letter and completion certificate. Mention and tag @Codveda in your posts.
- Use hashtags like #CodvedaJourney, #CodvedaExperience, and #FutureWithCodveda to showcase your progress and experiences.
- Share your project completion updates on LinkedIn, accompanied by a video explanation and the GitHub project repository link.
- You will be provided with four tasks. Select and complete any three tasks within your domain to fulfill the internship requirements.
- Submit your completed tasks via the Codveda submission form. Ensure all tasks are submitted within the allocated 15-day period.

SUBMISSION

- Create a professional video showcasing your internship projects and achievements.
- Host the video on LinkedIn to provide proof of your work and establish credibility among your peers. Consider tagging Codveda Technology in your posts to ensure they are notified of your work using hashtags like #CodvedaAchievements and #CodvedaProjects.
- A SUBMISSION FORM will be shared later. Till then, please continue your tasks and maintain a separate file for each level.
- When posting the video on LinkedIn, include engaging content that highlights your contributions and skills. Tailor the post to your specific internship domain to maximize impact and visibility.

Level 1 (Basic)



Task 1: Data Cleaning and Preprocessing

- Description: Work with a raw dataset (e.g., CSV file) that contains missing values, duplicates, and inconsistent data formats.

Objectives:

- Load the dataset using pandas.
- Identify and handle missing values (e.g., imputation or removal).
- Remove duplicate rows and standardize inconsistent data formats (e.g., date formats, categorical variables).

Tools: Python, pandas.

Level 1 (Basic)

Task 2: Exploratory Data Analysis (EDA)

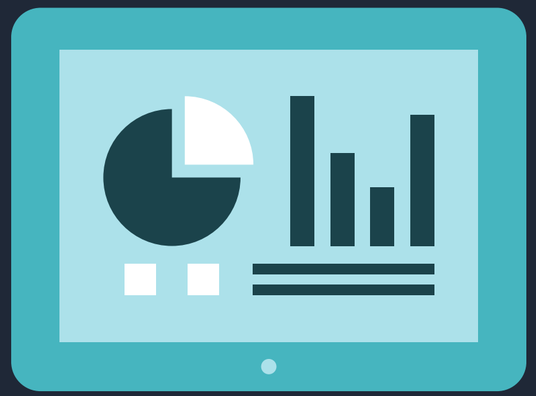
- Description: Perform an exploratory analysis on a given dataset to identify patterns, trends, and summary statistics.

Objectives:

- Calculate summary statistics (mean, median, mode, standard deviation).
- Visualize data distributions using histograms, boxplots, and scatter plots.
- Find correlations between numerical features.

Tools: Python, pandas, matplotlib, seaborn.

Level 1 (Basic)



Task 3: Basic Data Visualization

- Description: Create basic plots and charts to visualize the distribution and relationships within the dataset.

Objectives:

- Create bar plots, line charts, and scatter plots.
- Customize plot labels, titles, and legends.
- Export plots as images for reports.

Tools: Python, matplotlib, seaborn.

Level 2 (Intermediate)

Task 1: Regression Analysis

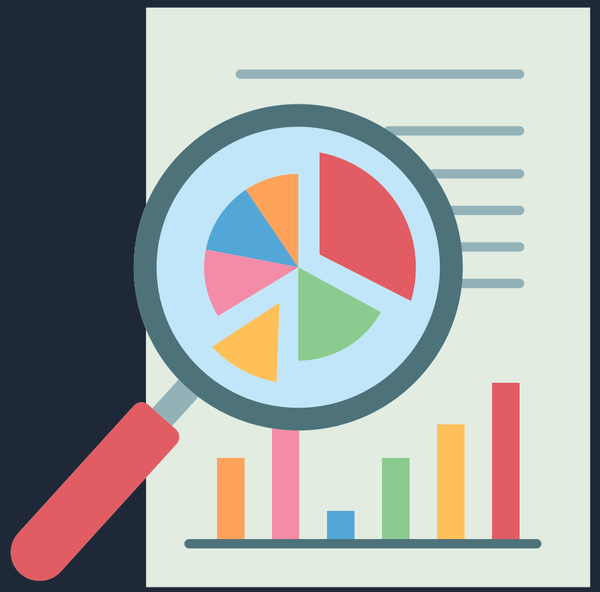
- Description: Perform a simple linear regression analysis to predict one variable based on another.

Objectives:

- Split the dataset into training and testing sets.
- Fit a linear regression model using scikit-learn.
- Interpret the coefficients and evaluate the model using metrics such as R-squared and mean squared error.

Tools: Python, scikit-learn, pandas

Level 2 (Intermediate)



Task 2: Time Series Analysis

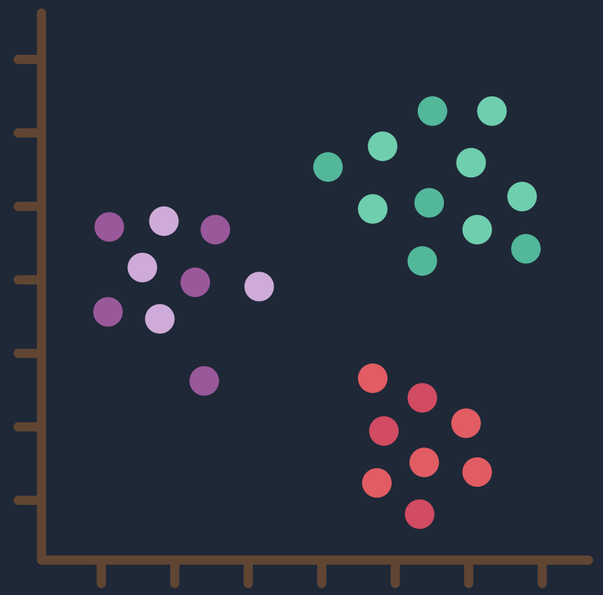
- Description: Analyze a time-series dataset (e.g., stock prices, temperature data) to detect trends and seasonality.

Objectives:

- Plot time-series data and identify patterns.
- Decompose the series into trend, seasonality, and residuals using statsmodels.
- Perform moving average smoothing and plot the results.

Tools: Python, pandas, matplotlib, statsmodels.

Level 2 (Intermediate)



Task 3: Clustering Analysis (K-Means)

- Description: Implement K-Means clustering to group similar data points together based on feature similarities.

Objectives:

- Standardize the dataset (e.g., using StandardScaler).
- Apply K-Means clustering and determine the optimal number of clusters using the elbow method.
- Visualize clusters using 2D scatter plots.
- Tools: Python, scikit-learn, matplotlib, seaborn.

Level 3 (Advanced)

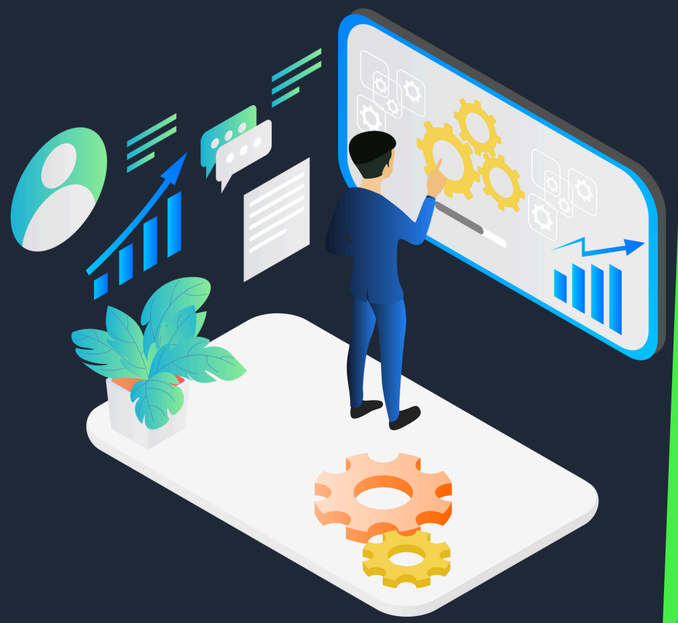
Task 1: Predictive Modeling (Classification)

- Description: Build and evaluate a classification model to predict categorical outcomes (e.g., predict if a customer will churn).

Objectives:

- Preprocess the data (handle categorical variables, feature scaling).
- Train and test multiple classification models (e.g., Decision Trees, Logistic Regression, Random Forest).
- Evaluate models using accuracy, precision, recall, and F1-score.
- Perform hyperparameter tuning using grid search.
- Tools: Python, scikit-learn, pandas, matplotlib.

Level 3 (Advanced)



Task 2: Building Dashboards with Power BI/Tableau

- Description: Create interactive dashboards using a data visualization tool like Power BI or Tableau to present insights from a dataset.

Objectives:

- Import and clean the dataset in Power BI/Tableau.
- Create interactive visualizations (e.g., bar charts, line graphs, maps).
- Set up filters and slicers for interactive exploration.
- Publish the dashboard and share it with others.
- Tools: Power BI or Tableau.

Level 3 (Advanced)

Task 3: Natural Language Processing (NLP) - Sentiment Analysis

- Description: Perform sentiment analysis on textual data (e.g., customer reviews, social media comments) to classify text as positive, negative, or neutral.

Objectives:

- Preprocess text data (tokenization, removing stopwords, and stemming/lemmatization).
- Use nltk or TextBlob for sentiment analysis.
- Visualize the sentiment distribution and word frequencies using word clouds.
- Tools: Python, nltk, TextBlob, pandas, matplotlib.

How to Contact Us?

For additional information, kindly get in touch with our team.

@codveda

support@technofyz.com

www.codveda.com