



INTERN  
INTELLIGENCE

# DATA ANALYTICS

Intern Intelligence

# ABOUT US

- Intern Intelligence is a pioneering organization dedicated to nurturing interns and achieving ambitious goals. Our mission is to support interns in building successful careers through innovative approaches and unwavering support. We aim to make significant strides and set new benchmarks in the industry, fostering the next generation of leaders.



# INSTRUCTION

- ◆ • **ID:** Your unique ID, provided in the offer letter, is crucial. Keep it safe as you will need it for task submission.
- ◆ • **Task Submission Link:** The link for task submission will be emailed to you within approximately one week.

# SUBMISSION

- ◆ • **Task Submission:** You will need to complete at least 2 tasks for successful completion of internship.
- ◆ • **Task Completion:** As part of your internship, you will be assigned several tasks. After completing each task, please record a video demonstrating your work and share it on LinkedIn, again using the hashtag **#internintelligence** and tagging **@InternIntelligence**.
- ◆ • **GitHub Repository:** Upload all completed tasks to GitHub. Name your repository in the format **InternIntelligence\_ProjectName**.

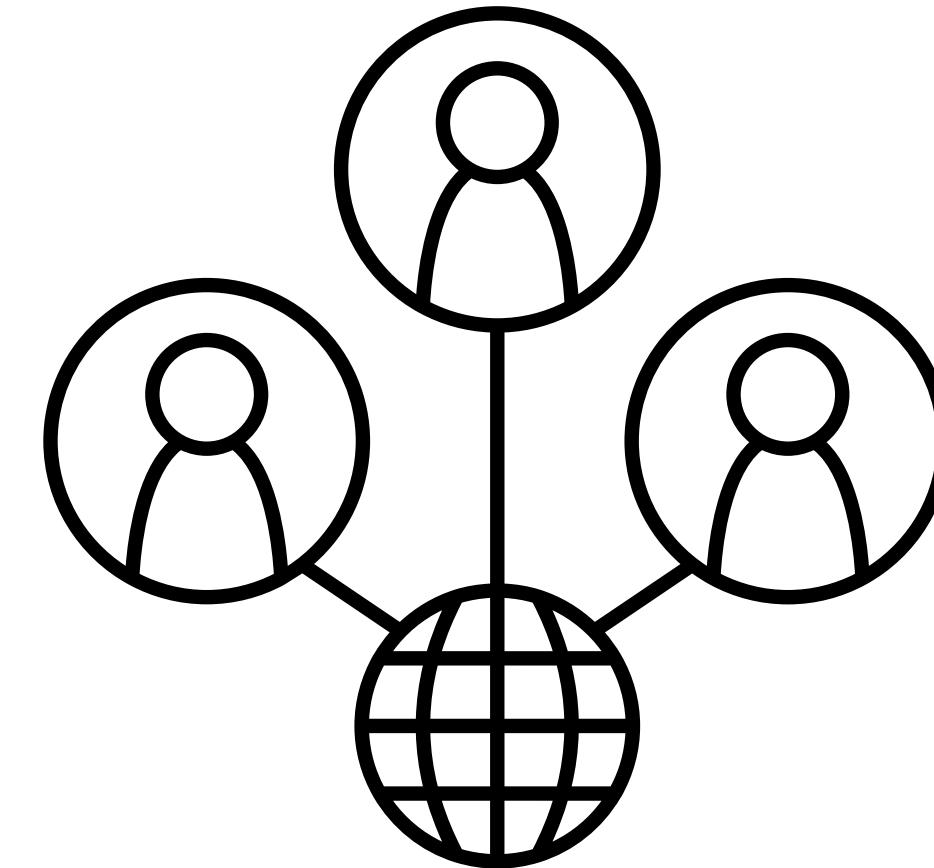
# About Internship



**COMPLETION  
CERTIFICATE**



**PLACEMENT  
SUPPORT**



**NETWORK  
OPPORTUNITY**

# DATA ANALYTICS TASK LIST



***You will need to complete at least 2 tasks for successful completion of internship.***

# TASK 1

## Predictive Analytics Model

**Task :** Develop a predictive analytics model to forecast future trends based on historical data.

### Details:

- **Data Preparation:** Clean and preprocess historical data.
- **Model Selection:** Choose appropriate machine learning models (e.g., regression models, time series forecasting).
- **Model Training:** Train the model using historical data and validate its performance.
- **Reporting:** Document the model's accuracy and provide insights based on the predictions.

### Where to Do It:

- **Jupyter Notebook:** Use for developing and testing machine learning models.
- **Google Colab:** An online platform for building and training models.
- **Kaggle:** Participate in competitions or use datasets for practice.

## TASK 2

# Data Visualization Dashboard

**Task :** Create an interactive data visualization dashboard to represent key metrics and trends.

### Details:

- **Data Integration:** Aggregate data from multiple sources.
- **Visualization Tools:** Use tools like Tableau, Power BI, or D3.js to create visualizations.
- **Interactivity:** Incorporate interactive elements such as filters and drill-down options.
- **Reporting:** Provide insights and recommendations based on the visualized data.

### Where to Do It:

- **Tableau:** Build interactive dashboards and visualizations.
- **Power BI:** Create and share data reports and dashboards.
- **D3.js:** Develop custom visualizations for web applications.

## TASK 3

# Advanced Statistical Analysis

**Task :** Perform advanced statistical analysis to uncover hidden patterns and relationships in the data.

### Details:

- **Statistical Techniques:** Apply techniques such as hypothesis testing, regression analysis, and factor analysis.
- **Data Exploration:** Explore data distributions, correlations, and anomalies.
- **Documentation:** Summarize findings and provide statistical insights and recommendations.

### Where to Do It:

- **RStudio:** Use R for statistical analysis and data exploration.
- **Python (SciPy, StatsModels):** Perform statistical analysis with Python libraries.
- **SPSS:** Conduct complex statistical analysis and modeling.

## TASK 4

# Big Data Analysis

**Task :** Analyze large datasets using big data tools and techniques to extract meaningful insights.

### Details:

- **Data Processing:** Use tools like Apache Hadoop or Spark to process large datasets.
- **Analysis:** Apply distributed computing techniques to perform data analysis.
- **Insights:** Derive insights and present findings from the big data analysis.

### Where to Do It:

- **Apache Hadoop:** For distributed storage and processing of large data sets.
- **Apache Spark:** Perform fast data processing and analysis.
- **Google BigQuery:** Analyze large datasets with SQL-like queries.

# TASK 5

## Data Pipeline Development

**Task :** Design and implement a data pipeline to automate data ingestion, transformation, and storage.

**Details:**

- **Data Sources:** Integrate various data sources and establish data pipelines.
- **ETL Process:** Develop Extract, Transform, Load (ETL) processes for data processing.
- **Automation:** Automate data workflows and ensure data quality.
- **Documentation:** Document the data pipeline architecture and processes.

**Where to Do It:**

- **Apache Airflow:** Orchestrate and automate workflows.
- **AWS Glue:** Develop and manage ETL workflows on AWS.
- **Google Dataflow:** Implement data processing pipelines in the cloud.

# Connect with us.



E-mail

[hr.internintelligence@gmail.com](mailto:hr.internintelligence@gmail.com)



LinkedIn

[@Intern Intelligence](https://www.linkedin.com/company/intern-intelligence)



Instagram

[InternIntelligence](https://www.instagram.com/internintelligence/)



Telegram

[InternIntelligence](https://t.me/InternIntelligence)



INTERN  
INTELLIGENCE