Data Science Practical Assessment - Intermediate Level

Total Time: 3 hours Maximum Points: 100

Instructions:

- Use Python for all solutions
- You may use pandas, numpy, sklearn, matplotlib, and seaborn libraries
- Show all your code and explain your approach for each solution
- Data files will be provided separately

Questions:

1. Customer Churn Analysis (5 points)

You are provided with `customer_data.csv` containing:

- customer_id
- subscription_length (months)
- monthly_charges
- total_charges
- internet_service
- contract_type
- churn (Yes/No)

Tasks: a) Clean the dataset and handle any missing values b) Create a predictive model for customer churn c) Identify the top 3 factors contributing to churn

2. Time Series Sales Forecasting (5 points) Using `sales_data.csv` containing daily sales for the past 2 years:

- date
- sales_amount
- promotion_active
- weekday
- holiday_flag

Tasks: a) Create a 30-day sales forecast using appropriate time series methods b) Analyze how promotions impact sales