

# Python Summer Party Challenge

by Interview Master

Day 5 of 15

## Nintendo

You are a Product Analyst working with the Nintendo Switch 2 pre-sales team to analyze regional pre-order patterns and customer segmentation. Your team needs to understand how different demographics influence pre-sale volumes across regions. You will leverage historical pre-sale transaction data to extract meaningful insights that can guide marketing strategies.

### Challenge Questions

**Q1:**

What percentage of records have missing values in at least one column? Handle the missing values, so that we have a cleaned dataset to work with.

**Q2:**

Using the cleaned data, calculate the total pre-sale orders per month for each region and demographic group.

**Q3:**

Predict the total pre-sales quantity for each region for September 2024. Assume that growth rate from August to September, is the same as the growth rate from July to August in each region.



# Want to try this yourself?

## Join the Challenge

Sign up for the Python Summer Party Challenge and solve 21 days  
of data science problems

[www.interviewmaster.ai/python-party](http://www.interviewmaster.ai/python-party)

Or keep scrolling to see my solutions

# My Solution - Q1

Day 5 Python Challenge

```
# Calculate percentage of records with missing values
missing_rows_percentage = pre_sale_data.isnull().any(axis=1).sum() / len(pre_sale_data) * 100
print(f"missing_rows_percentage:{.2f}")
```

  

```
# Handling missing values
cleaned_sale_data = pre_sale_data.dropna(subset=['pre_order_date', 'pre_order_quantity'])
cleaned_sale_data = pre_sale_data.fillna({
    'region': 'Unknown',
    'demographic_group': 'Unknown',
})
print(cleaned_sale_data)
```



# My Solution - Q2

Day 5 Python Challenge

```
# Get pre-order date month
pre_order_months = cleaned_sale_data['pre_order_date'].dt.month
cleaned_sale_data['pre_order_month'] = pre_order_months

# Calculate the total pre-sale orders per month for each region and demographic group.
total_pre_sale_orders_agg = (
    cleaned_sale_data.groupby(['region', 'demographic_group', 'pre_order_month'])
    .agg(total_pre_sale_orders=('pre_order_quantity', 'sum'))
    .reset_index()
)
print(total_pre_sale_orders_agg)
```



# My Solution - Q3

Day 5 Python Challenge

```
# Filter for July and August 2024
filtered_data = cleaned_sale_data[
    (cleaned_sale_data['pre_order_date'].dt.year == 2024) &
    (cleaned_sale_data['pre_order_date'].dt.month.isin([7,
8]))]
]

# Aggregate monthly sales for July and August by region
monthly_region_sales = (
    filtered_data.groupby(['region', 'pre_order_month'])['pre_
order_quantity']
    .sum()
    .unstack(fill_value=0)
    .rename(columns={7: 'July', 8: 'August'})
    .reset_index()
)

# Calculate the growth rate between July and August
monthly_region_sales['growth_rate'] = (
    (monthly_region_sales['August'] - monthly_region_sales['Ju
ly']) /
    monthly_region_sales['July'].replace(0, pd.NA)
)

# Predict September's sales
monthly_region_sales['September_prediction'] = (
    monthly_region_sales['August'] * (1 + monthly_region_sales
['growth_rate'])
).fillna(0).round(0).astype(int)

predicted_september_sales = monthly_region_sales[['region', 'S
eptember_prediction']]
print(predicted_september_sales)
```

Ready for your own challenge?

Try this yourself by signing up for the Python Summer Party challenge:

[www.interviewmaster.ai/python-party](http://www.interviewmaster.ai/python-party)

