

The screenshot displays a web browser window with multiple tabs open, including 'Data Analysis with Pandas', 'The Andro...', 'ChatGPT', 'Game Mo...', 'ChatGPT', 'Top A...', 'Shat...', and '1000'. The active tab is 'projects.datacamp.com/projects/619'. The browser's address bar shows the URL. Below the browser, a sidebar on the left contains a 'Project Instructions' section with a list of tasks: 'Task 10 Instructions', 'Load the user review data and plot it to visualize sentiment of paid vs. free apps.', 'Read datasets/user_reviews.csv into reviews_df DataFrame.', 'Merge apps and reviews_df DataFrames and assign the result to merged_df.', 'Create a box plot with Type on the x-axis and Sentiment_polarity on the y-axis.', 'If you'd like to learn more about sentiment analysis, check out DataCamp's Natural Language Processing Fundamentals in Python course.', 'Helpful links: pandas nerg() function documentation, boxplot() documentation.', 'Take Hint', and 'Some tests failed'. The main area shows a Jupyter notebook titled 'Project: The Android App Market on Google Play' with a 'notebook' tab selected. The notebook content includes a title '10. Sentiment analysis of user reviews', a paragraph about using review data to determine product quality, and a code cell with the following Python code:

```
In [34]: # Load user_reviews.csv
reviews_df = ...

# Join the two dataframes
merged_df = ...

# Drop NA values from Sentiment and Review columns
merged_df = merged_df.dropna(subset = ['Sentiment', 'Review'])

sns.set_style('ticks')
fig, ax = plt.subplots()
fig.set_size_inches(10, 8)

# User review sentiment polarity for paid vs. free apps
ax = sns.boxplot(x = 'Type', y = 'Sentiment', data = merged_df)
ax.set_title('Sentiment Polarity Distribution')

AttributeError: Traceback (most recent call last):
<ipython-input-34-7b1389180a2b> in <module>
      6
      7 # Drop NA values from Sentiment and Review columns
----> 8 merged_df = merged_df.dropna(subset = ['Sentiment', 'Review'])
      9 sns.set_style('ticks')
     10
AttributeError: 'AllData' object has no attribute 'dropna'
```

Load the user review data and plot it to visualize sentiment of paid vs. free apps:

- ### Code:

```
# Load the user reviews data
reviews_df = pd.read_csv('datasets/user_reviews.csv')

# Merge the DataFrames
merged_df = pd.merge(apps, reviews_df, on='App', how='inner')

# Drop NA values from Sentiment and Review columns
merged_df = merged_df.dropna(subset=['Sentiment', 'Review'])
```

```
# Set seaborn style
sns.set_style('ticks')
plt.figure(figsize=(10, 6))

# Create a box plot
sns.boxplot(x='Type', y='Sentiment_Polarity', data=merged_df)
plt.title('Sentiment Polarity Distribution')
plt.show()
```

Explanation:

1. `pd.read_csv()` is used to load the user review data from a CSV file into a DataFrame called `reviews_df`.
2. `pd.merge()` combines `apps` and `reviews_df` DataFrames on the 'App' column using an inner join, creating `merged_df`.
3. `dropna()` removes rows where 'Sentiment' or 'Review' columns have missing values.
4. `sns.set_style('ticks')` sets the visual style of the plots to 'ticks' for better readability.
5. A new figure is created with `plt.figure(figsize=(10, 6))` to set the size of the plot.
6. `sns.boxplot()` creates a box plot with 'Type' (Paid/Free) on the x-axis and 'Sentiment_Polarity' on the y-axis, using the `merged_df` DataFrame.
7. The title of the plot is set using `plt.title()`.
8. `plt.show()` displays the plot.