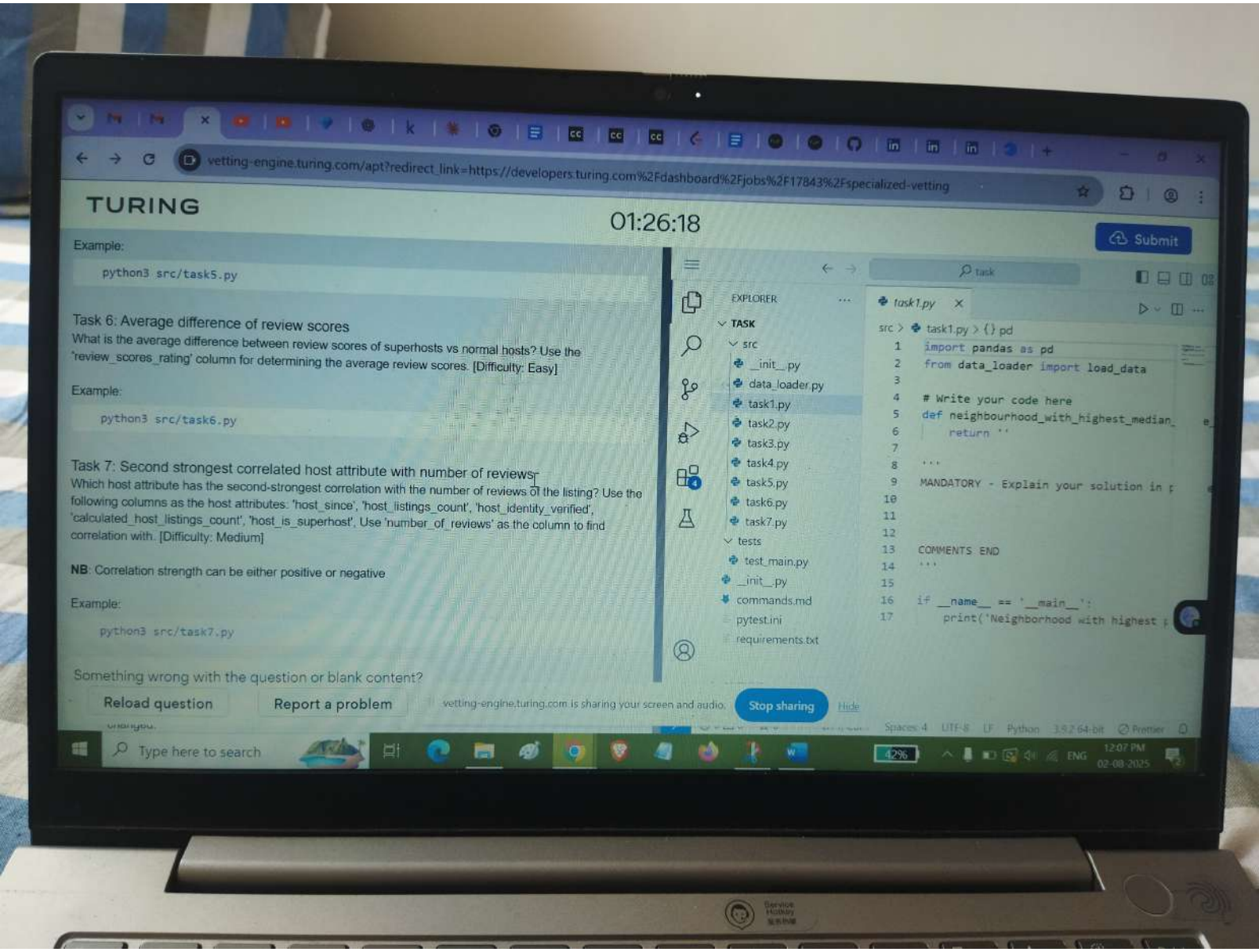
vetting-engine.turing.com/apt?redirect_link=https://developers.turing.com%2Fdashboard%2Fjobs%2F17843%2Fspecialized-vetting
TURING
01:26:18
Submit
Example:
python3 src/task5.py
Task 6: Average difference of review scores
What is the average difference between review scores of superhosts vs normal hosts? Use the 'review_scores_rating' column for determining the average review scores. [Difficulty: Easy]
Example:
python3 src/task6.py
Task 7: Second strongest correlated host attribute with number of reviews
Which host attribute has the second-strongest correlation with the number of reviews of the listing? Use the following columns as the host attributes: 'host_since', 'host_listings_count', 'host_identity_verified', 'calculated_host_listings_count', 'host_is_superhost'. Use 'number_of_reviews' as the column to find correlation with. [Difficulty: Medium]
NB: Correlation strength can be either positive or negative
Example:
python3 src/task7.py
Something wrong with the question or blank content?
Reload question
Report a problem
vetting-engine.turing.com is sharing your screen and audio.
Stop sharing
Hide
EXPLORER
TASK
src
__init__.py
data_loader.py
task1.py
task2.py
task3.py
task4.py
task5.py
task6.py
task7.py
tests
test_main.py
__init__.py
commands.md
pytest.ini
requirements.txt
task1.py
import pandas as pd
from data_loader import load_data
# Write your code here
def neighbourhood_with_highest_median_
return ''
'''
MANDATORY - Explain your solution in
COMMENTS END
'''
if __name__ == '__main__':
print('Neighborhood with highest
Type here to search
42%
ENG
12:07 PM
02-08-2025

vetting-engine.turing.com/apt?redirect_link=https://developers.turing.com%2Fdashboard%2Fjobs%2F17

TURING

01:26:04

Vacation Rental Listings and Reviews

Instructions

- Use the shortcut Ctrl + ` to open a terminal.
- Install required packages

```
pip install -r requirements.txt
```

- The commands on this page cannot be copy/pasted. However, they have been reproduced in `commands.md`. Feel free to copy them and paste to terminal for execution.
- The commands in commands.md should be run in the terminal and NOT PASTED in any files.
- You should only modify tasks related files. (E.g: task<taskNumber>.py)
- If you are familiar with Python notebooks, you can create .ipynb files for easier execution of the code, but make sure to remove before you submit the code.
- Do not modify any files with "# IMPORTANT ! DO NOT MODIFY THIS FILE" comment

Running/Testing the Code

You can run/test your code using either of the following 2 approaches

- run the unit tests in tests/test_main.py using the command:

```
pytest tests/test_main.py
```

- or execute each task file individually, the response would be logged to the console. The command to run each task file from the terminal is provided in the "Problem Statement"

Something wrong with the question or blank content? ...t cases in `tests/test_main.py` before

Reload question Report a problem

vetting-engine.turing.com is sharing your screen and audio. Stop sharing

EXPLORER

TASK

src

__init__.py

data_loader.

task1.py

task2.py

task3.py

task4.py

task5.py

task6.py

task7.py

tests

test_main.py

__init__.py

commands.md

pytest.ini

requirements.txt

- or execute each task file individually, the response would be logged to the console. The command to run each task file from the terminal is provided in the "Problem Statement"

We strongly encourage you to test your code using the provided test cases in `tests/test_main.py` before you submit.

Problem Statement

You are provided with a dataset containing vacation rental listings and reviews. Your task is to complete the code in seven separate Python task files, each addressing a specific question using the provided dataset.

Dataset description:

This dataset contains two files:

- **Listings**: This is a table that contains information about the vacation rental listings in a city. Every listing has a single row in the table and contains several attributes related to the listing as well as the host for the listing.
- **Reviews**: This is a table that contains information about all the guest reviews for the listings in the **"Listings"** table, along with related attributes such as the date of the review and the reviewer name.

Task 1: Neighborhood Price Difference

What is the neighborhood in which superhosts have the biggest median price difference with respect to non superhosts? Use the following three columns in the 'listings' dataset to answer this question: 'host_is_superhost', 'neighbourhood_cleansed', and 'price'. [Difficulty: Medium]

Example:

Something wrong with the question or blank content?

Reload question Report a problem

...has the strongest correlation to price? Use the following review score columns

TURING

01:25:38

```
python3 src/task1.py
```

Task 2: Correlation Analysis

Which of the review scores has the strongest correlation to price? Use the following review score columns in the 'listings' dataset: 'review_scores_rating', 'review_scores_accuracy', 'review_scores_cleanliness', 'review_scores_checkin', 'review_scores_communication', 'review_scores_location', 'review_scores_value'. [Difficulty: Easy]

NB: Correlation strength can be either positive or negative

Example:

```
python3 src/task2.py
```

Task 3: Professional Host Analysis

What is the average price difference between a professional host and a non-professional one? Consider a host as professional if they have listings in more than 5 different locations (location is defined by the 'neighbourhood_cleansed' column). [Difficulty: Medium]

Example:

```
python3 src/task3.py
```

Something wrong with the question or blank content?

Reload question Report a problem

entire homes / entire apartments and other types of listings

TURING

01:25:26

Task 4: Median Price Premium

What is the median price premium given to entire homes / entire apartments with respect to other listings of the same neighborhood? Report the average across all neighborhoods. Use the 'room_type' column in the 'listings' dataset to distinguish between entire homes / entire apartments and other types of listings. [Difficulty: Medium]

Example:

```
python3 src/task4.py
```

Task 5: Best Listing for Revenue

What is the listing with the best expected revenue based on the last 12 months, considering 60% of guests leave reviews and every guest will stay only the minimum number of nights? Use both the 'listings' and 'reviews' datasets for this question and only use listings with minimum nights of stay <= 7. The 'minimum_nights' column indicates the required minimum number of nights of stay for any listing. [Difficulty: Medium]

Example:

```
python3 src/task5.py
```

Task 6: Average difference of review scores

's vs normal hosts? Use the

Something wrong with the question or blank content?

Reload question Report a problem

vetting-engine.turing.com is sharing your screen and audio. Stop

Type here to search