DataInterview.com

60 Python Coding Interview Questions

*Seen in **Data Analyst**, **Data Scientist** and **Data Engineer** Interviews at FAANGs and startups*

Basic Python

1. What is the difference between a list and a tuple?
2. How do you create a lambda function?
3. What is the difference between deepcopy and shallowcopy?
4. Describe list comprehensions and give an example.
5. What does the enumerate() function do?
6. How do you handle exceptions in Python?
7. What is the difference between global and local variables?
8. How can you convert a string representation of a number into an integer?
9. Describe the use of the *args and **kwargs parameters.
10. What is the purpose of the __init__ method in a class?

String Manipulation

1. How do you reverse a string in Python?
2. Write a function to determine if two strings are anagrams.
3. How would you capitalize the first letter of each word in a string?
4. Write a function to find the longest common prefix among a list of strings.
5. How can you split a string by whitespace characters?
6. Describe the difference between str.replace() and str.translate().
7. Write a function to check if a string contains only digits.
8. How can you count the number of vowels in a string?
9. Write a function to convert a string into a list of words.
10. Describe the use of str.join() method with an example.

Algorithms

1. Describe the differences between a stack and a queue.
2. How can you implement a priority queue in Python?
3. Write a function to traverse a binary tree in in-order fashion.
4. Describe the characteristics of a hash table.
5. How do you determine if a linked list has a cycle?
6. Explain the difference between breadth-first search (BFS) and depth-first search (DFS).
7. Write a function to find the height of a binary tree.
8. How can you implement a graph data structure in Python?
9. Describe the difference between an array and a linked list.
10. Write a function to balance the brackets in a mathematical expression.
11. Explain the quicksort algorithm.
12. Describe the concept of dynamic programming with an example.
13. Write a function to determine if a number is prime.
14. Explain the Dijkstra's algorithm for finding the shortest path in a graph.
15. How do you find the kth smallest element in an unsorted array?
16. Describe the divide and conquer approach with an example.
17. Write a function to determine the longest increasing subsequence in an array.
18. Explain the difference between greedy algorithms and dynamic programming.
19. How do you implement the bubble sort algorithm?
20. Write a function to find the intersection point of two linked lists.

Pandas SQL Question at amazon

```
# Orders.csv
+----------+---------+------------+--------+
| order_id | user_id | order_date | amount |
+----------+---------+------------+--------+
|        1 |     101 | 2023-05-01 |  100.0 |
|        2 |     102 | 2023-05-03 |   50.0 |
|        3 |     101 | 2023-05-07 |  150.0 |
|        4 |     103 | 2023-05-10 |  200.0 |
|        5 |     102 | 2023-05-12 |   80.0 |
+----------+---------+------------+--------+

# Users.csv
+---------+------------+----------+
| user_id | join_date  | location |
+---------+------------+----------+
|     101 | 2023-01-01 | Seattle  |
|     102 | 2023-02-15 | Portland |
|     103 | 2023-04-05 | San Jose |
+---------+------------+----------+
```

1. Find the total amount spent by each user on their orders.
2. Identify users who have placed more than one order. Provide their user_id, location, and total number of orders placed.
3. For each location, find the average order amount and the latest order date.
4. Find the user_id of individuals who have made an order amounting to more than 100 but haven't made any orders after May 7, 2023.

Pandas SQL Question at Meta

```
# Posts.csv
+--------+---------+------------+---------+
| post_id| user_id | post_date  | content |
+--------+---------+------------+---------+
|      1 |     201 | 2023-03-15 | Photo   |
|      2 |     202 | 2023-03-16 | Text    |
|      3 |     201 | 2023-03-17 | Video   |
|      4 |     203 | 2023-03-18 | Photo   |
|      5 |     202 | 2023-03-19 | Link    |
+--------+---------+------------+---------+

# Reactions.csv
+----------+--------+---------+----------+
| react_id | post_id| user_id | reaction |
+----------+--------+---------+----------+
|        1 |      1 |     202 | Like     |
|        2 |      2 |     203 | Love     |
|        3 |      3 |     201 | Wow      |
|        4 |      4 |     201 | Haha     |
|        5 |      1 |     203 | Angry    |
+----------+--------+---------+----------+
```

1. Find the posts that have received reactions from their own authors. List the post_id and reaction.
2. For each type of reaction (Like, Love, etc.), calculate the total number of occurrences.
3. Find users who have posted content but have never reacted to any post, including their own.
4. For each user, identify the total number of reactions they've received on their posts. If they haven't received any reactions, they should still appear in the result with a count of 0.

Pandas SQL Question at Google

```
# Trips.csv
+--------+----------+-----------+-----------+
| trip_id| vehicle_id | start_time| end_time   |
+--------+----------+-----------+-----------+
|      1 |       501 | 2023-05-01 | 2023-05-01 |
|      2 |       502 | 2023-05-02 | 2023-05-02 |
|      3 |       501 | 2023-05-03 | 2023-05-03 |
|      4 |       503 | 2023-05-04 | 2023-05-04 |
|      5 |       502 | 2023-05-05 | 2023-05-05 |
+--------+----------+-----------+-----------+

# Sensors.csv
+----------+--------+------------+--------+---------+
| sensor_id| trip_id| sensor_type | status | readings|
+----------+--------+------------+--------+---------+
|        1 |      1 | Lidar       | Good   |   3521  |
|        2 |      2 | Camera      | Fault  |     0   |
|        3 |      3 | GPS         | Good   |   9872  |
|        4 |      4 | Lidar       | Good   |   3612  |
|        5 |      2 | GPS         | Good   |   9654  |
+----------+--------+------------+--------+---------+
```

1. Find all trips where at least one sensor reported a fault. List the trip_id and the corresponding sensor_type.
2. For each vehicle, calculate the average readings for each sensor type when the status is "Good". List the vehicle_id, sensor_type, and average readings.
3. Identify the vehicle with the highest number of trips. Provide the vehicle_id and the total number of trips.
4. Which sensor type has the most recorded data across all trips (based on readings)? Provide the sensor_type and the accumulated readings.

Pandas SQL Question at

AppStore.csv

app_id	app_title	genre	price
1	iChat	Social	0.00
2	iPhotoMagic	Photo	2.99
3	iTuneUp	Music	0.00
4	iFixTools	Utility	4.99
5	iRead	Books	0.00

Purchases.csv

purchase_id	app_id	user_id	purchase_date
1	1	401	2023-04-10
2	2	402	2023-04-11
3	1	403	2023-04-12
4	3	401	2023-04-13
5	2	401	2023-04-14

1. Identify the most expensive app in the App Store.
   Provide the app_title, genre, and price.
2. Calculate the total earnings from the Utility genre.
3. Which users have spent the most money on apps?
   Provide the user_id and the total amount spent.
4. Determine the genre popularity based on purchases.
   Provide the genre and its corresponding number of purchases, sorted in descending order of popularity.

Pandas SQL Question at NETFLIX

Shows

show_id	title	genre	seasons
1	Streamer's Life	Comedy	3
2	The Lost Byte	Thriller	1
3	Data Love	Romance	2
4	Query's End	Sci-Fi	4
5	Join Junction	Drama	2

Views

view_id	show_id	user_id	view_date	season
1	1	501	2023-05-01	1
2	2	502	2023-05-02	1
3	1	503	2023-05-03	2
4	3	501	2023-05-04	1
5	2	501	2023-05-05	1

1. Which show has the highest number of unique viewers? Provide the title and the count of viewers.
2. Calculate the total views per genre.
3. Identify users who have watched all the seasons of a particular show.
4. Find out which day had the highest streaming activity. Provide the view_date and the number of views for that day.