## Questions for Data Analyst Interview -

1. How to check if items in list A present in list B ?

Ans: Convert list to set and perform intersection on both sets

1. Convert the date string column to time series ?

Ans: pd.to\_datetime(date)

1. What do you mean by data integrity ?

Ans: Data integrity refers to the accuracy and consistency of data stored in database or data warehouse.

1. What are different types of clauses in sql ?

Ans: i) UPDATE, ii) GROUP BY, iii) FROM, iv) HAVING, v) ORDER BY, vi) USING, vii) WHERE

1. What is a Pivot Table ?

Ans: A PivotTable is an interactive way to quickly summarize large amounts of data.

1. What is a subquery in SQL ?

Ans: In SQL, placing a SELECT query inside another SELECT query is known as subquery.

1. Difference between Where and Having in sql ?

Ans: WHERE cannot be used with aggregates(sum, mean, etc) but HAVING clause can be used.

1. What is Data Wrangling ?

Ans: Data wrangling is the process of cleaning and integrating messy and complex data sets for easy access and analysis.

1. How to handle the missing data ?

Ans: Imputing the missing value, can replace the missing values with mean, median, forward fill and backward fill

1. Separate the given numeric values from column of the dataframe? (df[‘col1] = [‘a\_1’,’b\_2’,’c\_3’,’d\_4’]

Ans: import pandas as pd

df = pd.DataFrame(['a\_1','b\_2', ‘c\_3’, ‘d\_4’], columns = ['col1'])

df  = df.col1.str.split('\_')

1. What are Outliers?

Ans: An Outlier is a data point that differs significantly from other observations.

1. What are KPIs?

Ans: KPIs also known as Key performance indicators are the key targets you should track to make the most impact on your strategic business outcomes.

1. Write an SQL query to show records from one table that another table does not have.

Ans: Assume Id as primary key in both tables (table A, table B)

SELECT \* FROM tableA

WHERE id NOT IN (SELECT id FROM tableB)

1. Write a query to display the names of people in column my\_name\_list that begin with ‘H’.

Ans: SELECT my\_name\_list FROM table

WHERE my\_name\_list LIKE ‘H%’

1. Differentiate between Delete and Truncate.

Ans: The delete statement is used to remove single or multiple records from an existing table depending on the specified condition. The truncate command removes the complete data from an existing table but not the table itself.

1. What is Cross Join?

Ans: Cross Joins combines each row from the first table with each row from the second table

1. Count the Occurance of ‘name’ in given list using python [my\_list = [‘name’,’class’,’school’,’name’,’name’]

Ans:

count = 0

for i in l:

if i == 'name':

count +=1

print(count)

1. From given list, give the output my\_list[-1][-2] ( [my\_list = [‘name’,’class’,’school’,’name’,’name’])

Ans: ‘m’

1. Answer the following questions:
   1. Output for list1\*3 where list1 = [1,2]
   2. Output for list(‘cognitensor’)

Ans: a) [1, 2, 1, 2, 1, 2]

b) ['c', 'o', 'g', 'n', 'i', 't', 'e', 'n', 's', 'o', 'r']

1. What this # symbol do in python?

Ans: # symbol is used to write comments in the python code