

1. Enumerate a few SQL commands

SQL

- SELECT ... FROM ... JOIN .. ON .. WHERE ... GROUP BY ...
HAVING ... ORDER BY ... LIMIT
- CREATE TABLE
- avg(), min(), max(), cast(), substr()
- INSERT
- TRUNCATE, CASE, JOIN ... ON
- ALTER TABLE .. ADD
- DROP TABLE
- AS
- UPDATE
- DELETE
- C.R.U.D. (create, read, update, delete)

psql commands

```
\copy  
\c  
\l  
\dt  
\h
```

terminal (bash) command

```
psql  
pg_dump
```

2. What is a Primary Key?

Unique, not null identifier of rows in a table. Used to find rows quickly.

3. What is a Foreign Key?

Reference to the Primary Key of another table. Creates a unidirectional connection between two tables. To see Foreign Keys, an Entity-Relationship diagram (ER diagram) is super useful.

4. Which 5 things you need to connect to a PostgreSQL DB?

- host (localhost, 127.0.0.1, IP address, AWS endpoint)
- port (5432) (22 SSH, 80 HTTP)
- username (default is your system user)
- password
- DB name (default is your system user)

Store passwords in environment variables:

add to .bashrc

```
export POSTGRES_PASSWORD=1234
```

update and check on bash:

```
source .bashrc
```

```
echo $POSTGRES_PASSWORD
```

access from Python

```
import os
```

```
os.getenv('POSTGRES_PASSWORD')
```

?? how to edit environment variables / .bashrc on Windows?

5. Name a few programs that connect to a remote machine

- psql
 - ssh
 - scp
 - browser
 - pg_dump
 - email
 - requests
 - pip
 - git push / pull
-

6. Compare pandas and PostgreSQL. What are pros and cons of one or the other?

Postgres:

- good for storing data
- good for big data
- faster to run

- strict data types
- access control
- version control (transactions)

Pandas:

- good for working on data
 - visualization
 - calculations, shorter path to ML
 - faster to write
-

7. Who is Slonik?

- the blue Postgres elephant