

```
INSERT INTO customer (SELECT * FROM
snowflake_sample_data.tpch_sf10.customer);
runtime: 5.46s
```

```
INSERT INTO supplier (SELECT * FROM
snowflake_sample_data.tpch_sf10.supplier);
runtime: 2.01s
```

```
INSERT INTO region (SELECT * FROM
snowflake_sample_data.tpch_sf10.region);
runtime: 1.25s
```

```
INSERT INTO partsupp (SELECT * FROM
snowflake_sample_data.tpch_sf10.partsupp);
runtime: 6.67s
```

```
INSERT INTO part (SELECT * FROM snowflake_sample_data.tpch_sf10.part);
runtime: 3.81s
```

```
INSERT INTO orders (SELECT * FROM
snowflake_sample_data.tpch_sf10.orders);
runtime: 10.02s
```

```
INSERT INTO nation (SELECT * FROM
snowflake_sample_data.tpch_sf10.nation);
runtime: .575s
```

```
INSERT INTO lineitem (SELECT * FROM
snowflake_sample_data.tpch_sf10.lineitem);
runtime: 30.44s
```

```
--QUERY #1
select min(supplier.s_name), count(*) as total_parts
from supplier, partsupp
where supplier.s_suppkey = partsupp.ps_suppkey
group by supplier.s_suppkey;
```

```
Runtimes (in ms):
74,51,63,47,73,80,54,82,202,83,52,90,49,46,49,56,79,72,60,53
```

```
Results (first two rows):
Supplier#000070034      80
Supplier#000045032      80
```

```
Number of Rows: 100,000
```

```
--QUERY #2
```

```
select max(ps_supplycost)
from partsupp;
```

Runtimes (in ms):

48,63,311,208,42,40,52,68,619,247,185,857,43,97,44,47,51,47,47,44

Result: 1000.00

Number of Rows: 1

--QUERY #3

```
select max(partsupp.ps_supplycost)
from supplier, partsupp
where supplier.p_suppkey = partsupp.ps_suppkey
group by supplier.p_suppkey;
```

Runtimes (in ms):

103,54,86,51,62,100,336,163,53,608,73,57,61,49,50,378,54,53,50,49

Results (first two rows):

975.99

994.39

Number of Rows: 100,000

--QUERY #4

```
select min(nation.n_name), count(*) as total_customers
from nation, customer
where nation.n_nationkey = customer.c_nationkey
group by nation.n_nationkey;
```

Runtimes (in ms):

69,86,179,54,109,104,303,375,57,64,70,54,63,88,170,59,68,56,52,49

Results (first two rows):

UNITED KINGDOM 60381

GERMANY 60153

Number of Rows: 25

--QUERY #5

```
select min(x.s_name), count(*) as num_orders
from supplier x, lineitem y
where x.s_suppkey = y.l_suppkey
      and ( y.l_shipdate between '1996-10-10' and '1996-11-10')
group by x.s_suppkey;
```

Runtimes (in ms):

68,127,81,342,317,778,77,51,132,536,53,76,57,47,54,58,139,415,48,49

Result (first two rows):

Supplier#000016137	5
Supplier#000099031	9

Number of Rows: 99,964

Discussion of Runtimes

The insertion queries took significantly longer than the queries in Part 4. Within the multiple runtimes of the queries in Part 4, they tended to have the same quick runtime, but occasionally there would be a runtime much longer than the others. This is why the standard deviation is so high relative to the mean. In addition, I found it interesting that the second query, which seemed relatively simple because it did not require any joins, took longer on average than 3 out of the other 4 queries.

Mean and Standard Deviation of Query Runtime

