Johnathan Clementi
Database Systems CMPT 308
Professor Alan G. Labouseur
September 11th, 2017
Lab 2 Short Essays and Screenshots

1. Execute Basic queries:
    a. Select * from Customers;

```
147  select *
148  from Customers;
149
150  select *
```

Data Output | Explain | Messages | History

	cid character (4)	name text	city text	discountpct numeric (5,2)
1	c001	Tipt...	Dulu...	10.00
2	c002	Tyrell	Dallas	12.00
3	c003	Eldon	Dallas	8.00
4	c004	ACME	Dulu...	8.50
5	c005	Wey...	Risa	0.00
6	c006	ACME	Beiji...	0.00

    b. Select * from Agents;

```
150  select *
151  from Agents;
152
153  select *
154  from Products;
155
156  select *
```

Data Output | Explain | Messages | History

	aid character (3)	name text	city text	commission numeric (5,2)
1	a01	Smith	New...	5.60
2	a02	Jones	New...	6.00
3	a03	Perry	Hon...	7.00
4	a04	Gray	New...	6.00
5	a05	Otasi	Dulu...	5.00
6	a06	Smith	Dallas	5.00
7	a08	Bond	Lon...	7.07

c. Select * from Products;

```
152
153  select *
154  from Products;
155
156  select *
```

Data Output | Explain | Messages | History

	pid character (3)	name text	city text	qty integer	priceusd numeric (10,2)
1	p01	Heis...	Dallas	111400	0.50
2	p02	univ...	New...	203000	0.50
3	p03	Com...	Dulu...	150600	1.00
4	p04	LCA...	Dulu...	125300	1.00
5	p05	pencil	Dallas	221400	1.00
6	p06	trap...	Dallas	123100	2.00
7	p07	flux ...	New...	100500	1.00
8	p08	HAL...	New...	200600	1.25

d. Select * from Orders;

```
155
156  select *
157  from Orders;
```

Data Output | Explain | Messages | History

	ordno integer	month character (3)	cid character (4)	aid character (3)	pid character (3)	quantity integer	totalusd numeric (12,2)
1	1011	Jan	c001	a01	p01	1100	495.00
2	1012	Jan	c002	a03	p03	1200	1056.00
3	1015	Jan	c003	a03	p05	1000	920.00
4	1016	Jan	c006	a01	p01	1000	500.00
5	1017	Feb	c001	a06	p03	500	540.00
6	1018	Feb	c001	a03	p04	600	540.00
7	1019	Feb	c001	a02	p02	400	180.00
8	1020	Feb	c006	a03	p07	600	600.00
9	1021	Feb	c004	a06	p01	1000	457.50
10	1022	Mar	c001	a05	p06	450	810.00
11	1023	Mar	c001	a04	p05	500	450.00
12	1024	Mar	c006	a06	p01	880	400.00
13	1025	Apr	c001	a05	p07	888	799.20
14	1026	May	c002	a05	p03	808	711.04

2. A superkey is considered any combination of columns that can uniquely differentiate between rows of table. A candidate key is a "minimal superkey," meaning that it is a superkey with the smallest number of columns to have unique values for every row. Finally, the primary key is a value found in one column that can uniquely identify the row (record).

3. SQL rely on table schemas to dictate the data type needed for each column and their records. When creating the table schemas, the maximum length of each record for the column is defined when defining the column's datatype. SQL systems such as PostgreSQL support the data types CHAR, VARCHAR, BIT, BIT VARYING, BOOLEAN, INT, FLOAT, DECIMAL, DATE and TIME. The CHAR(n) and VARCHAR(n) data types are character strings with a defined length (n). BIT(n) and BIT VARYING(n) are the same as CHAR and VARCHAR except that instead of holding character strings, they hold strings of bits. BOOLEAN values are the same as other languages – TRUE, FALSE, or UNKNOWN. INT's are integers and FLOAT's are floating point numbers (same as other languages). DECIMAL (n, d) denotes a number of places for a whole number (n) and a number of places for its fractional decimal (d). Finally, DATE and TIME are a special form of character strings that hold dates and times. Below is a possible table with its columns and their datatypes defined.

   Concerts

cid	CHAR(4), not NULL
city	TEXT, nullable
artist	TEXT, not NULL
venue	TEXT, nullable
costUSD	Numeric (4,2), nullable

4. Explain the relational rules:
   a. The "first normal form rule" explains that data must be indivisible or atomic. In other words, data in columns should be in its most basic form. For instance, a name such as 'Johnathan Clementi' can be broken down into firstName: 'Johnathan' and lastName: 'Clementi'. This rule is important for the robustness and simplicity of databases because if it was not implemented, databases would be much more unwieldy to deal with.
   b. The "access row by content only" rule (also known as the "What, not where" rule explains that data must be queried based on data values itself and not the location of said data. This is important because data locations can change easily by adding or deleting from the table.
   c. Finally, the "all rows must be unique" rule explains that there should not be any duplicate data because it is just dumb… seriously, duplicate data doesn't make sense – it's just taking up space in memory that could be used for other things. It could also alter calculations in which the number of occurrences of something is

important (such as the work I did over the summer. If I had duplicate data of fish from the same lake, then my calculations of the overall sex ratio would be biased and invalid).