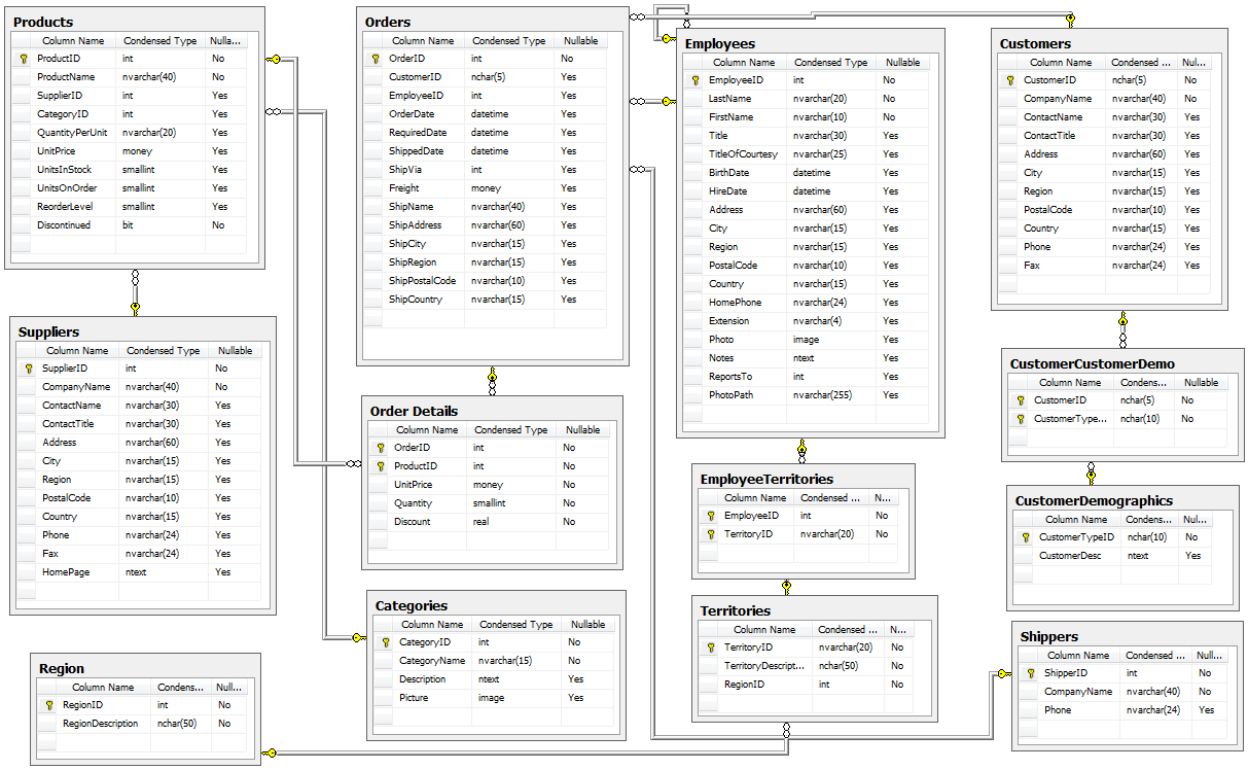
Econ 524 SQL Assignment

Download the northwind.db SQLite database. Here is the schema for the database. Note that the schema is not 100% accurate for the table and field names. In order to get your queries working, you may need to run some preliminary queries to find out the true names for tables/fields.



To get a list of all the tables, you can run the following query.

```
select name from sqlite_master where type='table';
```

Name your script northwind.py. There are 8 queries to run and one function to write. Save the results of the 8 queries in data frames named d1, d2, d3, d4, d5, d6, d7, and d8. Make sure that the columns in your data frames are in the same order than I list them in the instructions. Also, you should do all data transformation/manipulation in your SQL queries (i.e. you should not be using pandas or other python functions to help except for the read_sql_query function).

1. Create a dataframe with columns for all fields from the Orders table and that includes only orders that shipped to the USA.
2. Create a dataframe with one column, that contains the names of all countries where the company has customers. (Each country should only appear once in the DataFrame)
3. Create a dataframe with two columns: (1) the names of countries where the company has customers and (2) the number of customers from that country. Only include countries with more than one customer. Order your results so that countries with more customers appear first.
4. Create a dataframe with one column, that contains the OrderID for all orders that customers shipped to a different country than customer's country.

5. Create a dataframe with two columns: orderID and total order revenue. You can calculate the total order revenue from the ordedetails table. In the orderdetails table, there is one record for each product. The revenue from that product: (1-discount)*UnitPrice*Quantity. An order may contain more than one product and the total order revenue is the sum of the revenue for each product in the order.
6. Create a dataframe with three columns: orderId, order date, and total order revenue. Only include orders placed by a Customer based in the USA.
7. Create a dataframe with one column, the company name, for customers who have shipped at least one order to Eugene (the shipcity). Companies should only appear at most once.
8. Create a dataframe with one column, the company name, for customers who have shipped at least two orders to Eugene (the shipcity). Companies should only appear at most once.
9. Write a function named orderlookup that takes two arguments (both of type str): city and country. It returns a data frame with columns for each field from the orders table and that contains all records for orders in the provided city and country. (Note: your function should connect to the database to run the query, i.e. create a new connection object in your function).