Name: _____

Uniqname: _____

Question 1a. [5 points] Fill in the following code to create the Pandas dataframe below:

	A	B
0	1	2
1	3	4
2	5	6
3	7	8

```
df = pd.DataFrame(


)
df
```

Question 1b: [5 points] Write one line of code to create a data frame **from the one above** that looks like:

	A	B
0	8	7
1	6	5
2	4	3
3	2	1

Name: _____

Uniqname: _____

Question 2: [10 points] Fill in the [ ] after each instance of the word “In” with the numbers 1 through 8 to indicate the order in which the code should be arranged to produce the following output.

```
SEX                F         M      diff   absdiff
COURSE
PSYCH350    3.374767  3.027596  0.347170  0.347170
ECON101     2.424861  2.784909 -0.360047  0.360047
SPANISH101  3.185170  2.812277  0.372893  0.372893
BUDDHST230  3.626923  3.230909  0.396014  0.396014
CMPTRSC280  1.945783  2.349241 -0.403458  0.403458
```

Note that the ‘SUBJECT’, ‘CATALOG_NBR’, ‘GRD_PTS_PER_UNIT’ and ‘COURSE’ columns are contained in student.course.csv and ‘SEX’ is in student.record.csv. You may assume that both the studentRecord and studentCourse DataFrames are indexed on the same ID values.

In [ ]:
```
print(df.sort_values('absdiff').tail())
```

In [ ]:
```
df['absdiff'] = df['diff'].apply(abs)
```

In [ ]:
```
df = pd.merge(studentCourse,studentRecord,how='left')
```

In [ ]:
```
df = df.pivot_table(index=['COURSE'],
                    columns=['SEX'],
                    values='GRD_PTS_PER_UNIT',
                    aggfunc=np.mean)
```

In [ ]:
```
df['COURSE'] = df['SUBJECT'] + df['CATALOG_NBR'].apply(str)
```

In [ ]:
```
studentRecord = pd.read_csv('student.record.csv')
studentCourse = pd.read_csv('student.course.csv')
```

In [ ]:
```
import numpy as np
import pandas as pd
from scipy import stats
```

In [ ]:
```
df['diff'] = df['F'] - df['M']
```

Name: _____

Uniqname: _____

Question 5: [10 points (2 points per blank)] Given the following DataFrame,

	Account	Name	Rep	Manager	Product	Quantity	Price	Status
0	714466	Trantow-Barrows	Craig Booker	Debra Henley	CPU	1	30000	presented
1	714466	Trantow-Barrows	Craig Booker	Debra Henley	Software	1	10000	presented
2	714466	Trantow-Barrows	Craig Booker	Debra Henley	Maintenance	2	5000	pending
3	737550	Fritsch Russel and Anderson	Craig Booker	Debra Henley	CPU	1	35000	declined
4	146832	Kiehn-Spinka	Daniel Hilton	Debra Henley	CPU	2	65000	won
5	714466	YQG	Jane Booker	Fred Anderson	CPU	1	10000	presented
6	714466	YQG	Jane Booker	Fred Anderson	Software	1	10000	presented
7	714466	YQG	Jane Booker	Fred Anderson	Monitor	1	5000	pending

fill in the blanks of the following pivot_table function to give the pivot table shown below.

```
pd.pivot_table(df,
                index=[price],

                columns=[product],

                aggfunc=[sum],

                values=[price],

                fill_value= 0,

                dropna=True)
```

		sum			
		Price			
	Product	CPU	Maintenance	Monitor	Software
Manager	Status				
Debra Henley	declined	35000	0	0	0
	pending	0	5000	0	0
	presented	30000	0	0	10000
	won	65000	0	0	0
Fred Anderson	pending	0	0	5000	0
	presented	10000	0	0	10000