# Assignment: Creating, Updating and Using Index Assignment

***1. Use the code below to import the assignment data-set.***

Of course make sure that you ***first*** download the file (***brooklynsales.csv***) (**it is attached as a resource**) and then import it into whichever software you are using.

PS Don't forget to c***hange the location of file in the infile statement*** (as the below code is the location of where I have the file on my computer. )

data brooklysales;

infile "E:\Workspace\index\brooklynsales.csv" DSD MISSOVER FIRSTOBS=2;

input borough neighborhood$ buildingclass$ tax_class$ b_class$ address$ zip$ salesp$;

run;

2. Answer ***all*** the questions posed. ***Assume*** that it makes sense to make an index in this scenario.

## Questions for this assignment

1. Which *variables* in the data set would be poor candidates as index

variables, and *how* did you confirm this?

2. Create a *simple index, on the zip variable,* using *proc sql* for the brooklynsales *existing* data set.

3. How can you *confirm* that you successfully created your simple index from question 2? *Hint: What is the SAS code to find out?*

4. Use the code below now, and tell me if the simple index you created (*in question 2*) was *utilized?* If it was utilized, how do you know? If it wasn't utilized, how do you know?

```
data useindex; set brooklynsales; where zip=11201 or zip=11249; run;
```

5. Before you answer this question. *Here is a little background to help you out.*

There are *many reasons* why SAS might *not* utilize your Index, even if you created one.

Some reasons are *broad/general*.

An example of a *broad/general* reason is *poor selection of the discriminant variable.*

If you make a *poor selection* in terms of the discriminant variable, *it may actually take longer to use the index than not. So SAS decides not to use your Index.*

More *specific reasons* have to do with using *certain statements, expressions, or forms of statements/expressions*, that will ensure that your Index is not used. For example, did you know that *you can't* use an if statement if you want to exploit your Index?

*Question:*

With that in mind, *how can you change the where expression from question #4 so that your index is utilized?*
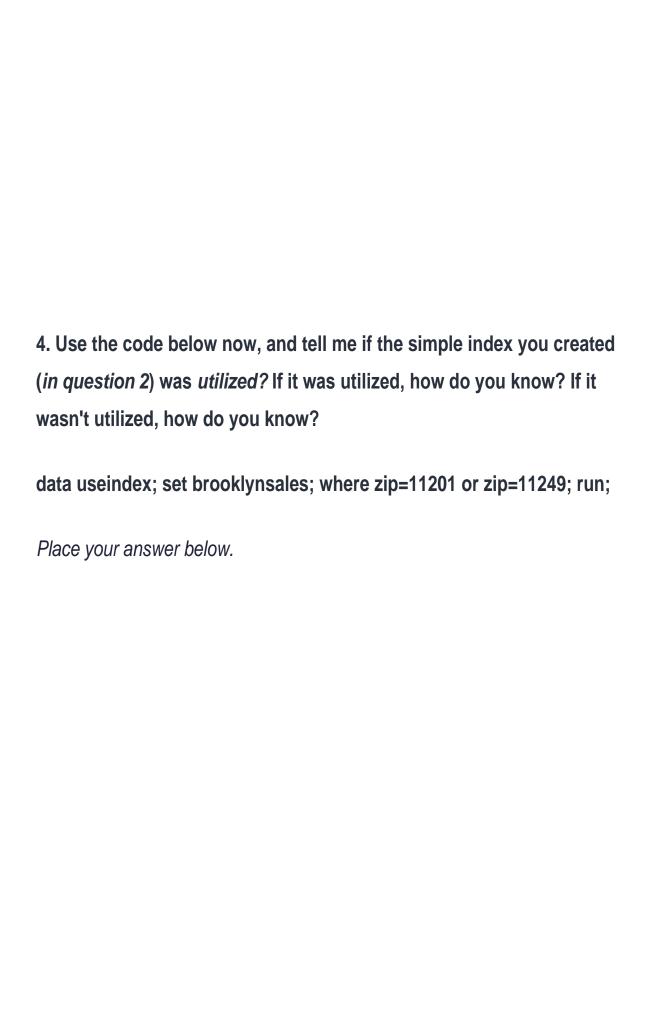
# *Download resource files below*

**Download**

Assignment submission

Save or submit your work

**1. Which *variables* in the data set would be poor candidates as index variables, and *how* did you confirm this?**

*Place your answer below.*

**2. Create a *simple index, on the zip variable,* using *proc sql* for the brooklynsales *existing* data set.**

**3. How can you *confirm* that you successfully created your simple index from question 2?** *Hint: What is the SAS code to find out?*

*Place your answer below.*

**4. Use the code below now, and tell me if the simple index you created (*in question 2*) was *utilized?* If it was utilized, how do you know? If it wasn't utilized, how do you know?**

**data useindex; set brooklynsales; where zip=11201 or zip=11249; run;**

*Place your answer below.*

**5. Before you answer this question.** *Here is a little background to help you out.*

There are *many reasons* why SAS might *not* utilize your Index, even if you created one.

Some reasons are *broad/general*.

An example of a *broad/general* reason is *poor selection of the discriminant variable.*

If you make a *poor selection* in terms of the discriminant variable, *it may actually take longer to use the index than not. So SAS decides not to use your Index.*

More *specific reasons* have to do with using *certain statements, expressions, or forms of statements/expressions*, that will ensure that your Index is not used. For example, did you know that *you can't* use an if statement if you want to exploit your Index?

*Question:*

With that in mind, *how can you change the where expression from question #4 so that your index is utilized?*

*Place your answer below.*

**6.Choose your sharing preference**

I want to get feedback from my fellow students (yes or no)?

*Place your answer below.*

1. Which *variables* in the data set would be poor candidates as index variables, and *how* did you confirm this?

2. Create a *simple index, on the zip variable,* using *proc sql* for the brooklynsales *existing* data set.

3. How can you *confirm* that you successfully created your simple index from question 2? *Hint: What is the SAS code to find out?*

4. Use the code below now, and tell me if the simple index you created (*in question 2*) was *utilized?* If it was utilized, how do you know? If it wasn't utilized, how do you know?

```
data useindex; set brooklynsales; where zip=11201 or zip=11249; run;
```

5. Before you answer this question. *Here is a little background to help you out.*

There are *many reasons* why SAS might *not* utilize your Index, even if you created one.

Some reasons are *broad/general*.

An example of a *broad/general* reason is *poor selection of the discriminant variable.*

If you make a *poor selection* in terms of the discriminant variable, *it may actually take longer to use the index than not. So SAS decides not to use your Index.*

More *specific reasons* have to do with using *certain statements, expressions, or forms of statements/expressions*, that will ensure that your Index is not used. For example, did you know that *you can't* use an if statement if you want to exploit your Index?

*Question:*

With that in mind, *how can you change the where expression from question #4 so that your index is utilized?*

# Give feedback to 3 other students

Reflecting on other students' work is likely to increase your own understanding