

Homework 2

Due: Sept. 9, 2010, 2:25 pm

Reminders

For the following problems you may discuss the concepts that will help solve these problems with classmates and course staff. You may *not* simply copy down the answers of your classmates as that is a violation of the collaboration policy. The one exception to this rule are those problems marked as Independent. You may discuss independent problems *with course staff only*.

Pivot tables

Throughout this course, we'll tend to work with very organized data. It'll be very common to have lots of instances that are very similar, which we can organize in columns. For example, suppose we have records for a class: each week we have a quiz or two, and each student gets a grade. (Equivalently, we have a senate, and every so often we have a vote; each senator casts a vote. Or we're running Netflix, and every so often a customer rents a movie.) Our data might look like this (for a very small class):

Name	Week	QuizScore
Amy	1	80
Mary	1	72
Robin	1	90
Amy	2	85
Robin	2	91
Mary	2	65
Mary	2	79
Robin	2	90
Amy	2	70
Mary	3	90
Robin	3	80
Amy	3	79

We've entered the grades into our spreadsheet in the order the papers happened to fall during grading, so some weeks Amy is first, other weeks Mary is first, etc. Notice that in week two, we had two quizzes rather than just one.

This spreadsheet is a kind of *database*, in the sense that it's a bunch of *records*, where each record provides information about a single name/week/quiz score. The key thing about these records is that each consists of the same three things: a name, a week, and a quiz score. These three things are called *fields* in the database. (We'll see databases quite a lot more later in the course, so this is just a nice easy example to get you used to some terminology.)

Now imagine that we'd like a summary of each student's average performance, by week. We want a table that looks like this:

	Week1	Week2	Week3
Amy	80	77.5	79
Mary	...		
Robin			

To do this, we can create what's called a "pivot table," and Excel has a tool that can produce this for us.

There is a tutorial on how to do pivot tables Excel [here](#).

Task 1:

Download `ExcelBasicsII.xlsx`, and do Task 1 in the spreadsheet, which asks you to create a pivot table using the students' weekly average.

A note about pivot tables

There's one tricky thing about pivot tables: they only work for summarizing numbers. If you had a table of letter grades, then things like "max" and "min" and "average" don't work, and pivot tables are of no use to you. You can still use them, however, by a little subterfuge: suppose you replaced each A with a 5, each B with a 4, and so on. Then you could compute the maximum grade, or the minimum. The "average" would not make sense. You could compute it, but it would have no real meaning. (Why not?) When you created the pivot table of maximums, however, you'd have a table full of 5s, 4s, 3s, etc. That's not ideal. So you could take that table and replace all the 5s with As, 4s with Bs, and so on. Summary:

- Convert letter grades to numerical representations
- Make pivot table based on numerical representations

- Convert pivot table results back to letter grades

There's one case where this convert-pivot-convert-back approach is especially effective: when each entry in the pivot table corresponds to exactly one row in the input table. For instance, in class we'll look at senators' votes on various bills. Each senator gets exactly one vote (Yea/Nay/Not Voting) on each bill. In this case, the operation that's done on each cell "max, min, average, ..." will always end up computing that single votes.

Concatenating two cells

You have already seen the addition of two cells in Excel, but sometimes you would like to combine the contents of two cells in other ways. If you have 42 in cell A1 and foo in cell A2, you can make foo:42 in cell A3 by typing = A1 & ":" & A2 into A3. The & operation combines two strings of characters by concatenating them.

Task 2:

On Sheet1 of ExcelBasicsII.xlsx in cell E26, make the contents read hello, world! as a combination of E23 and E24, spaces included (Task 2 on spreadsheet).

More If Statements

Task 3:

Final Exercise: (Independent) In cell F31, enter a formula that produces a 5 if E31 has an A in it, a 4 if E31 has a B in it, and so on. Hint: try this formula:

```
=If( E31 = "A", 5, If(E31 = "B", 4, 700))
```

Change it a bit so that it does work, and fill in the column labeled “Numerical Representation” with your formula. Does this work for all letter grades (A, B, C, D, F)? By the way, if the value in cell A1 is neither A, nor B, nor C, etc., what should you put in cell A2? We have to make a choice; a good idea is to put something that will stand out, like **Bad Data**. Do so. **Note:** The quotation marks around the A and the B and the **Bad Data** are necessary. If you don’t use them, Excel will think that A is the name of some cell (remember names from Homework 1?) The quotations marks tell Excel that this thing is to be treated explicitly as text, and *not* as a number or a name.

Final Handin

Task 4:

1. Rename your Excel file to something like JohnHughesHW2 (i.e., your name, followed by HW2).
2. Make sure all 4 tasks are completed.
3. E-mail your completed file to the TAs at cs0931tas@cs.brown.edu with the **Subject:** line HW2.