**VLOOKUP**

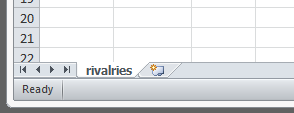
We’re going to start by looking at whether major rivalries between teams affect game statistics in systematic ways. For instance, are there more total penalty yards during a rivalry game than a non-rivalry game?

1. Start by downloading a table that contains a list of a the 9 top rivalries in the NFL (in the LAB EXERCISES folder in our Oncourse site, called Week04\_Rivalries.xlsx):

|  |  |  |
| --- | --- | --- |
| rivalry | team1 | team2 |
| CINvCLE | CIN | CLE |
| ATLvNO | ATL | NO |
| DALvWAS | DAL | WAS |
| KCvOAK | KC | OAK |
| CHIvGB | CHI | GB |
| BALvPIT | BAL | PIT |
| NYGvPHI | NYG | PHI |
| MIAvNYJ | MIA | NYJ |
| INDvNE | IND | NE |
|  |  |  |

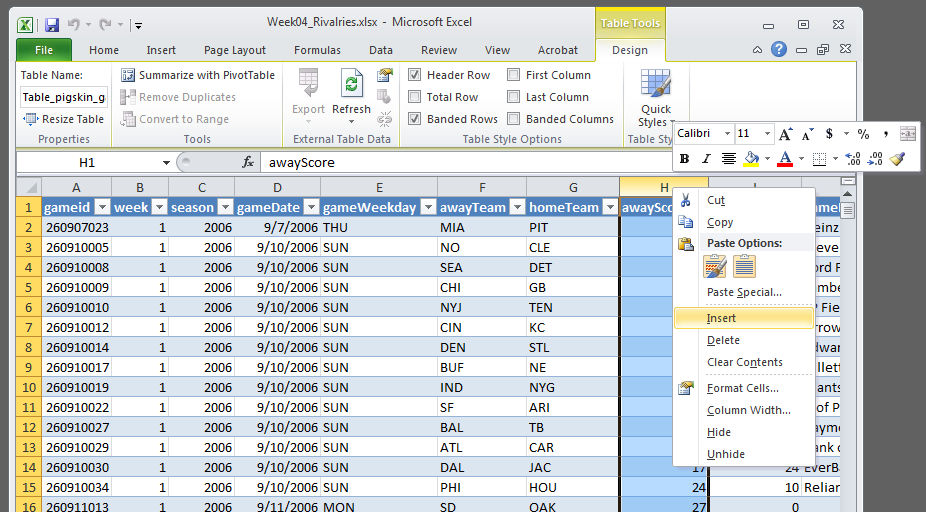
This table was intentionally designed so that teams only appear once in this list (despite the fact that some (e.g., NE, DAL, and PIT) are involved in multiple rivalries), and so that the teams appear in alphabetical order. (team1 < team2, alphabetically)

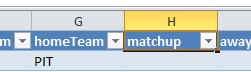
1. Open the rivalries workbook in Microsoft Excel, and create a new sheet by clicking the icon next to the rivalries tab:



This will create a new sheet called Sheet1

1. In Sheet1, access the **games** table from the pigskin database.
2. Insert a new column between the homeTeam column and the awayScore column, by right clicking on the header for column H, and selecting “Insert”:



1. Rename the new column to “matchup”  
   
2. We’re going to create a formula that will return a string that represents the two teams playing one another (in the same format as the “rivalries” column in the rivalries worksheet).
   1. The first step is to create a formula that will combine the two teams into one lookup\_value that we can search for in the rivalries table. In cell H2, type:  
      =CONCATENATE([@awayTeam],“v”,[@homeTeam])
   2. So far so good, but there’s a problem... Sometimes the CINvCLE rivalry will appear as CLEvCIN (whenever Cleveland is playing at Cincinnati), and Excel won’t find CLEvCIN in the rivalry table. **Try to write an IF statement that will put the homeTeam before the awayTeam in the rivalryFlag column only if the home team is before the away team alphabetically.**
3. Now we’re going to create a new column called “rivalriesFlag” just next to the matchup column that we just created:  
   
4. In this new **rivalryFlag** column, we want it to say “Rivalry” or “Not a Rivalry” depending on whether the matchup appears in the rivalries table. To do this, we need to use VLOOKUP.
   1. The first step is to create a formula that will look for the matchup value in the rivalries table. Write this formula in the rivalryFlag column, in cell I2:  
      =VLOOKUP([@matchup],rivalries!A$2:C$10,1,FALSE)

…which looks for the value of [@matchup] on the rivalries table between A2 and C10, and returns the value in the 1st column (rivalries) if it finds it, and #N/A otherwise.

* 1. So far so good, but we don’t just want #N/A errors if a matchup isn’t a rivalry. Let’s build on the formula that you wrote above, and use the ISNA function to detect whether the value in rivalryFlag is an error.  
     =ISNA(VLOOKUP([@matchup],rivalries!A$2:C$10,1,FALSE))  
     …which just takes the formula you’d written in the previous step, and puts it in an ISNA function.
  2. Now, if a matchup is a rivalry it will return FALSE, and it’ll return TRUE otherwise… but that’s not a good thing—we really want the opposite, for it to return TRUE if it’s a rivalry. **Try to use the =NOT() function to fix this problem.**
  3. Great! Now we want it to return “Rivalry” when the value is TRUE, and “Not a Rivalry” when the value is FALSE. **Write an IF statement that’ll do this.**

1. Now that you’ve added detail about rivalries, build a PivotTable, and start answering the questions on this week’s worksheet.

**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Week 6 Lab: VLOOKUP**

1. Using the “rivalryFlag” (described in the previous steps) fill-in the following values:

|  |  |  |
| --- | --- | --- |
|  | Rivalry | Not a Rivalry |
| Average number of turnovers for both teams |  |  |
| Average number of penalty yards for both teams |  |  |
| Average absolute value (=ABS()) of the difference between scores |  |  |
| Average gameAttendance |  |  |

(hint: For the first three, you might want to make new columns in the table)