Data Analysis Using Spreadsheets – crib sheet part 2

(Continued from crib sheet part 1...)

**Grouped aggregation**

Average temperature by country column F =unique(b:b), column G =averageif(b:b,f2,e:e)

**Your Turn** (1: 47.5, 2: 51.07, 3: 41.42, 4: 47.2)

2: =averageif(e:e,”<10”,c:c) 4: two filters + copy-paste, or =averageifs(c:c,a:a,“=*a”,b:b,”=*a”)

**Joining**

Show Countries sheet. Goal: find number of cities in the EU

Copy-paste Countries data into columns G,H,I,J

Add three columns between B and C

=index(k:k,match(b2,j:j), extend to column

=index(l:l,match(b2,j:j), extend to column

=index(m:m,match(b2,j:j), extend to column

Delete referenced data, show need to copy-paste using Paste special > values only

Finally use =countif(d:d,”=yes”) for number of cities in EU

*Note: match() function assumes sorted column; add third parameter 0 if not sorted*

**Pivot tables**

- Select entire table, then Data > Pivot table
- Average temperature by country: Rows country, Values temperature AVERAGE
- Average temperature for each coastline-EU combination: Rows coastline, Columns EU, Values temperature AVERAGE, then add Values population MAX; show copy-paste special to new sheet; alternative EU as second Rows instead of Columns
- Number of cities in different countries: Rows country, Values city COUNTA
- Note same result when add Values coastline COUNTA, switch to COUNTUNIQUE
- Back to city COUNTA, sort by: COUNTA of city

**Your Turn** (in the EU, Andorra, Moldova)

1) pivot table, Rows country, columns coastline, values longitude AVERAGE, sort by AVERAGE of longitude in no

2) pivot table, Rows country, filter coastline no, values longitude AVERAGE, sort by AVERAGE of longitude

3) pivot table, Rows coastline, Rows country, values longitude AVERAGE, in Rows country sort by AVERAGE of longitude