# Exploring Data with Pivot Tables

**Hypothesis:** In the districts with the most success in increasing government primary school student enrollment, the quality of education has been compromised.
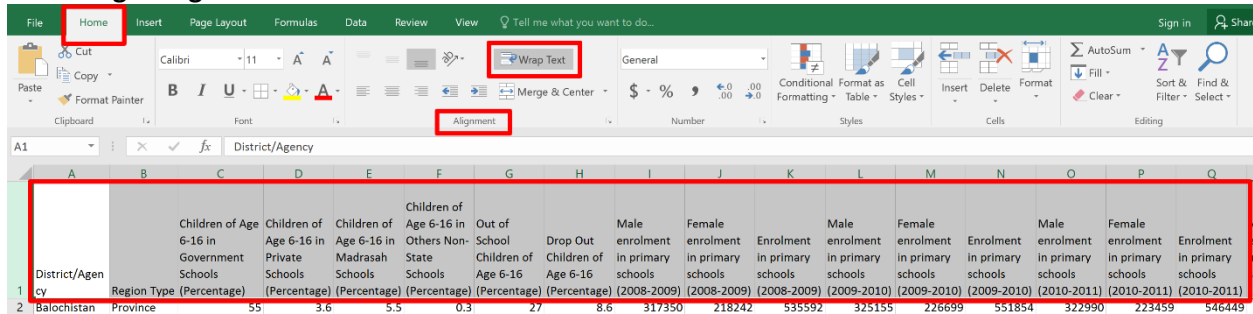
- Question 1 Which districts have the highest rate of Children of Age 6-16 in Government Schools (Percentage)?

- Question 2 Which districts experienced the greatest percentage change: Enrolment in primary schools (2008-2009) to Enrolment in primary schools (2010-2011) Average percentage change of enrollment in primary school?

- Question 3 What was the percentage change in Primary schools (2008-2009) and Sum of Primary schools (2010-2011) Average percentage change of primary school?

- Question 4 What was the percentage change in teaching staff from teaching staff in primary schools (2008-2009) teaching staff in primary schools (2010-2011) Average teaching staff change for all schools?

- Question 5 Which districts have the highest:
  Percentage of Children of class 5 who can read letters in urdu
  Percentage of Children of class 5 who can read words in urdu
  Percentage of Children of class 5 who can read sentences in urdu
  Percentage of Children of class 5 who can read story in urdu

- Question 6 Which districts have the highest:
  Percentage of children of class 5 who can recognize number from 1-9
  Percentage of children of class 5 who can recognize number from 10-99
  Percentage of children of class 5 who can subtract two digits
  Percentage of children of class 5 who can divide two digits

**Preparing the data**

1. Open **Balochistan Education Statistics 2014.xlsx**
2. There are two sheet tabs **Source** and **Balochistan**. In source tab sheet keep all data related with the source, helpful links, definitions of different terms in the dataset, and glossary of abbreviations. The Baluchistan tab is the original data.
3. **Make a copy:** Right click on Balochistan original sheet use the Move or Copy function, click on move to the end and create a copy.
4. **Rename the copy:** Rename Balochistan (2) Sheet as WORKING.
5. Save changes.

**Question 1 Which districts have the highest rate of Children of Age 6-16 in Government Schools (Percentage)?**

6.  **ALWAYS** review the dataset, make sure you understand all column headers and indicators and if you don't: look for them, google them, find the definitions in the document or webpage that you downloaded the data set from.
7.  One way to ensure you can read the headers when they are long is to use Wrap Text. Select Column 1 in Main Menu go to Home- Alignment – Wrap Text.
8.  **For working with PivotTables, all your columns must have headers in the first row with data beginning in the second row**.



9.  Delete Column A Country, B Province/Territory, and E Date. Select the columns and right click and select Delete. The data in these columns is identical or .
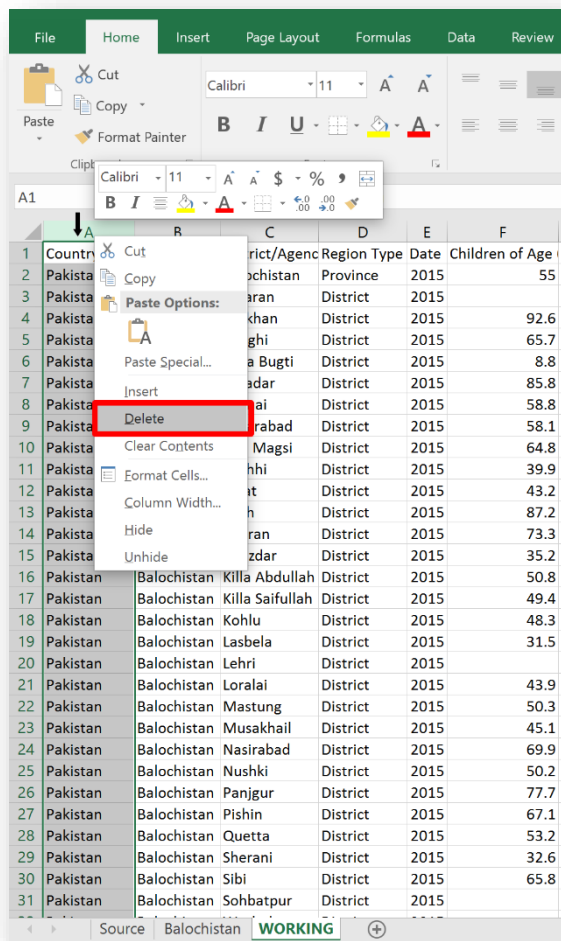
10.    As you can see Column B **Region Type** has just two variables, Province and Districts. The province is a sum. We will be calculating our own sums and we want each row to represent just one district. Delete row 2 with Balochistan data and Column B Region Type.

**Question 1 Which districts have the highest rate of Children of Age 6-16 in Government Schools (Percentage)?**

11.    Go to WORKING Sheet
12.    Select all data
13.    In the Main Menu go to Insert – Table – Pivot Table and Click OK in the Create PivotTable window
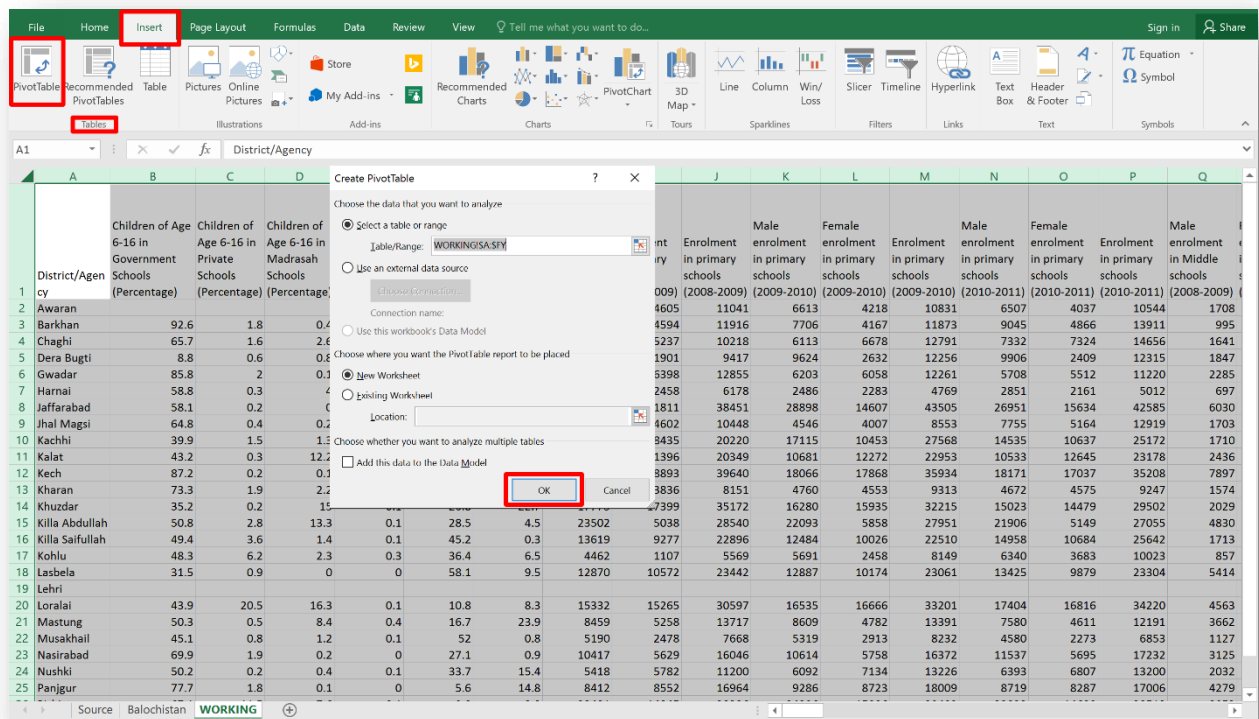
**14.** Rename the new Sheet as Q1

**15.** A cell inside the PivotTable has to be selected to be able to see the Field List (or the menu panel for PivotTables). In case you have lost it, select one cell inside the PivotTable or go to the Main Menu – PivotTable Tools – Analyze – Show Field List.
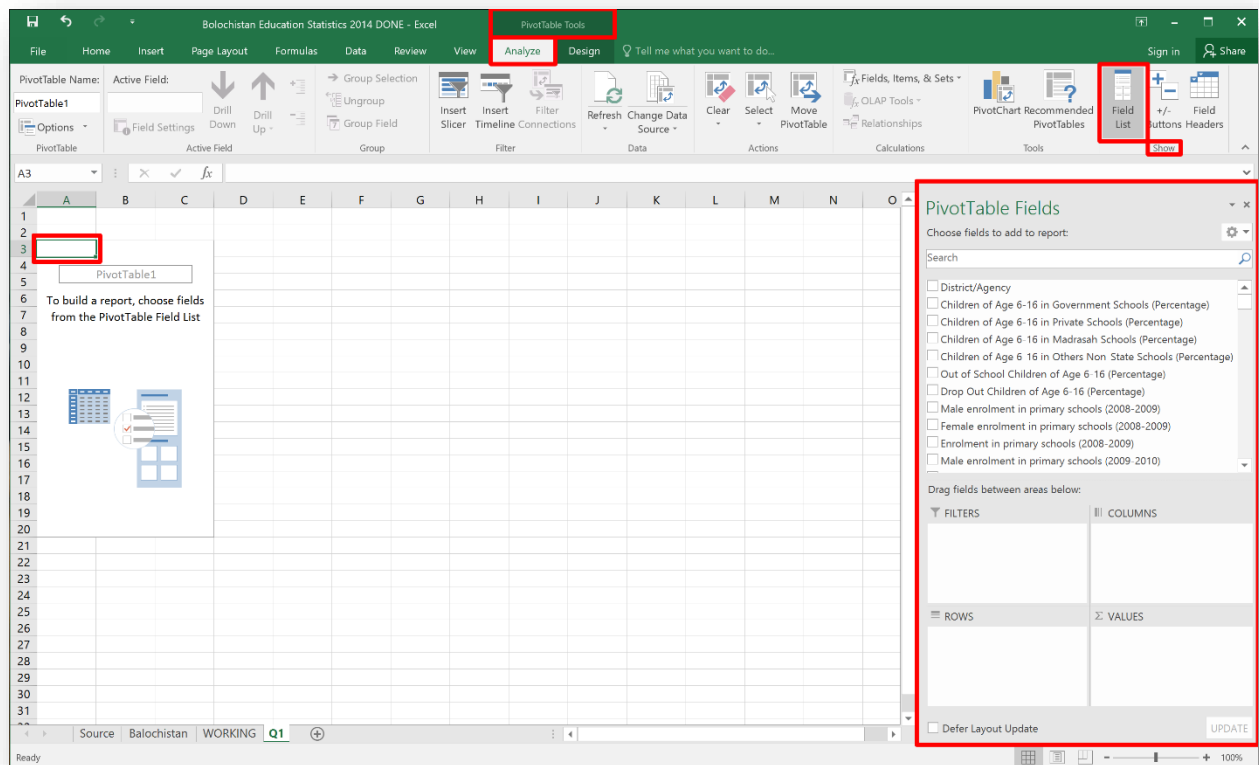
16.     **Pivot Table Fields** panel on the right has five main components:
   - The first field displays all the columns from the dataset: in this case WORKING sheet.
   - The second one is Filters, which allows you to filter the different fields
   - The third one is Columns, where you drag the field(s) than you want to display as distinct columns in the pivot table
   - The forth is Rows, where you drag the field(s) than you want to show in the row(s) in the pivot table
   - And Values is a tool for mostly numeric fields to do calculations.
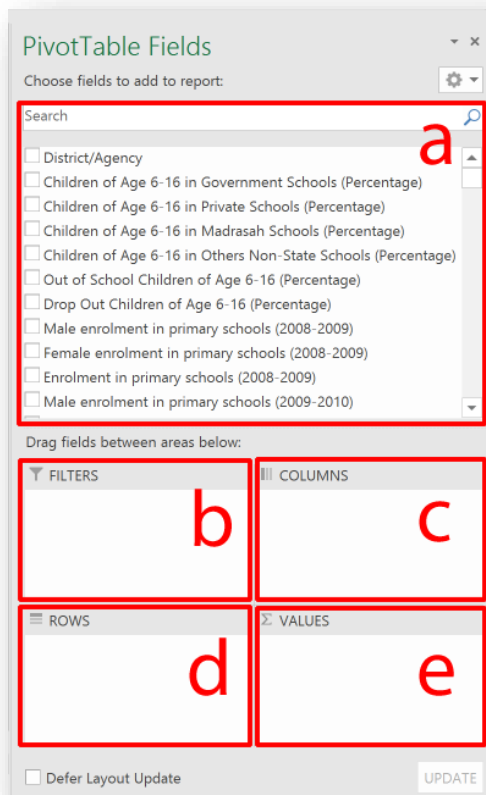
17. Ok, question 1 is about the column in WORKING sheet called Children of Age 6-16 in Government Schools (Percentage) and the highest rate by districts. Subsequently, we need the districts as a rows and Children of Age 6-16 in Government Schools (Percentage) as a Value in the PivotTable. Consequently:
    ● Click and drag District/Agency field in the Rows panel. You will see all the districts appear in column A in the PivotTable
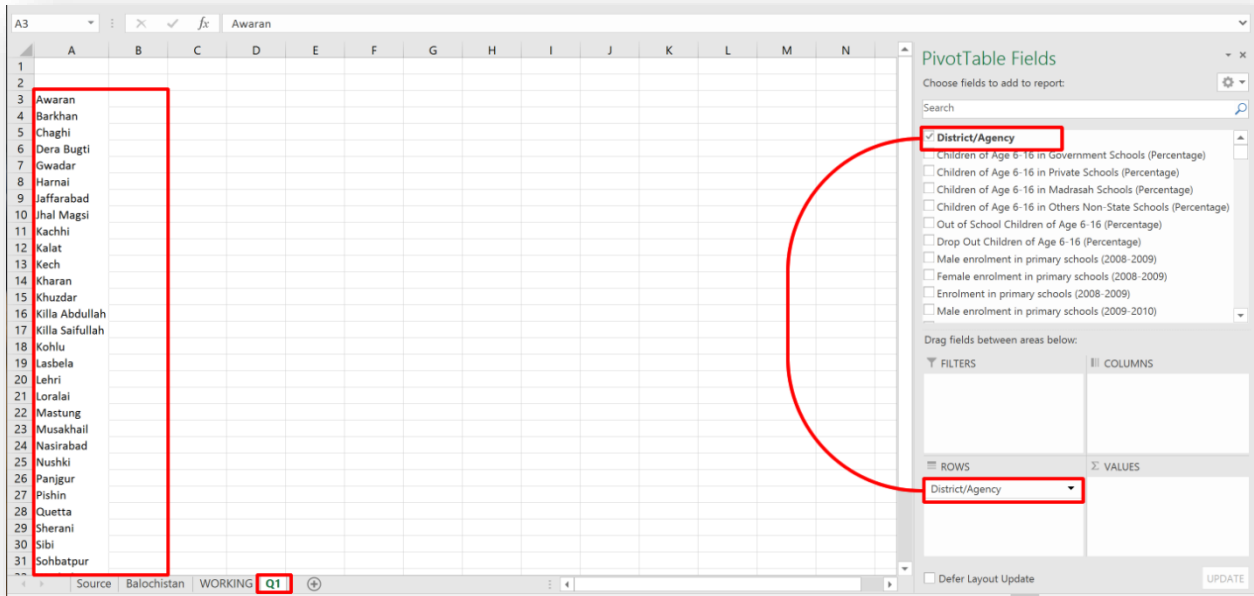
- Click and Drag Children of Age 6-16 in Government Schools (Percentage) field in Values panel.
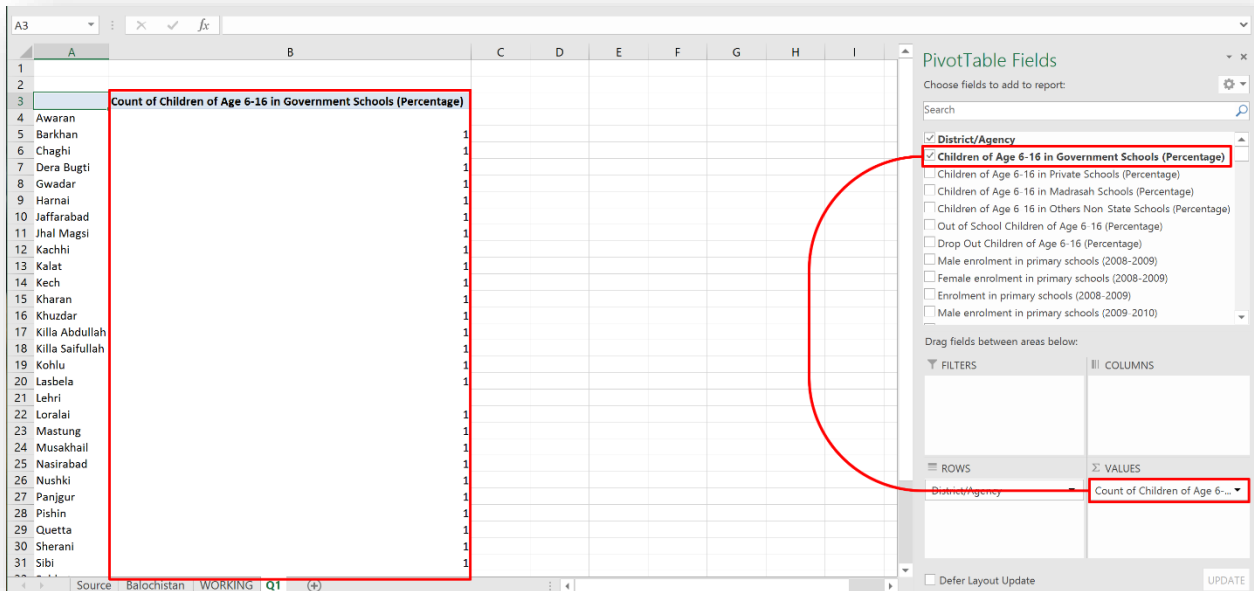


- There is a problem! In most cases by default PivotTables does a count of elements when you add numeric fields in Values panel. To change it, go to the small arrow inside Children of Age 6-16 in Government Schools (Percentage) field in the Values panel, click on Value Field Setting and in the Value Field Setting window click on Sum to add the values in each row and click on OK.
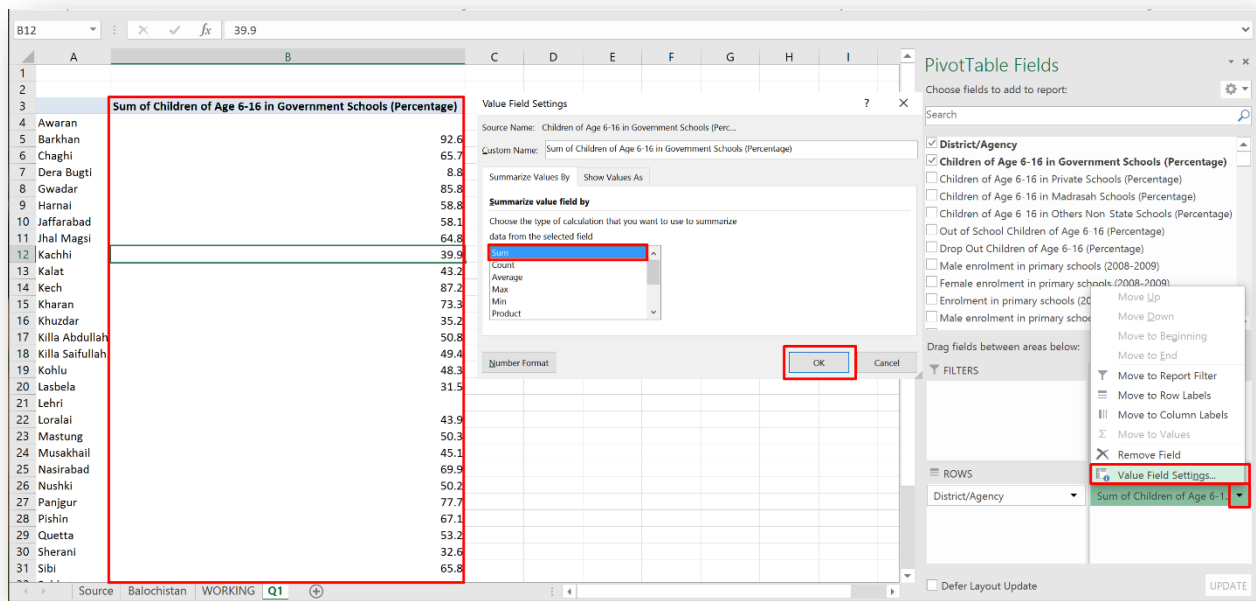
**18.**   Now we need the highest district at the top; you can sort the result in the Pilotable by right-clicking on any cell of Children of Age 6-16 in Government Schools (Percentage) column and in the option panel click on Sort – Sort Largest to Smallest.
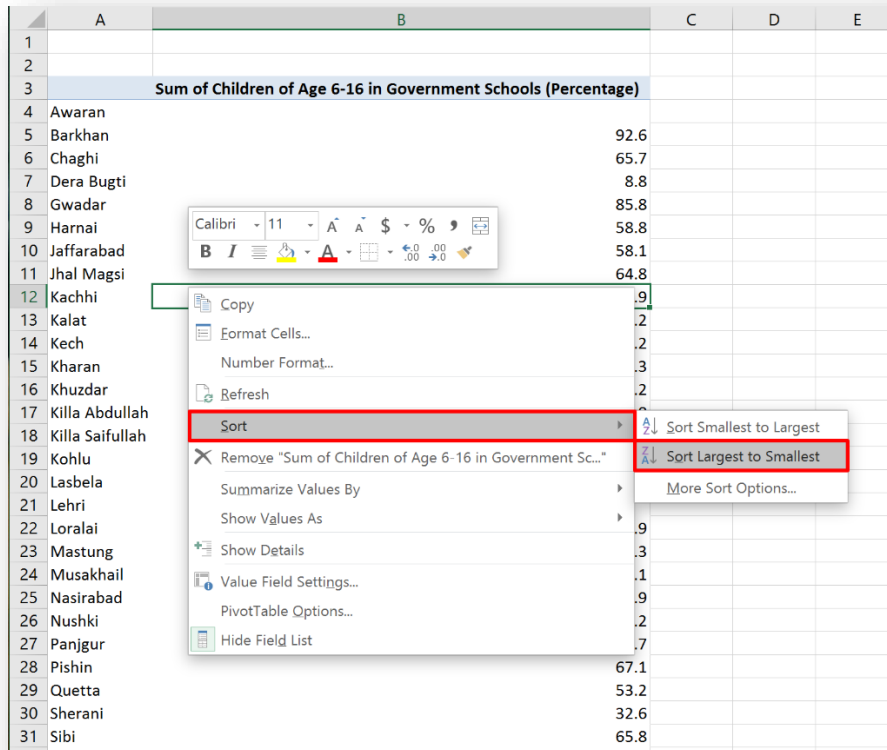
**19.**   Question 1 Which districts have the highest rate of Children of Age 6-16 in Government Schools (Percentage)?

|  | Sum of Children of Age 6-16 in Government Schools (Percentage) |
|---|---|
| Barkhan | 92.6 |
| Kech | 87.2 |
| Gwadar | 85.8 |
| Panjgur | 77.7 |
| Kharan | 73.3 |
| Nasirabad | 69.9 |
| Pishin | 67.1 |
| Sibi | 65.8 |
| Chaghi | 65.7 |
| Jhal Magsi | 64.8 |

**Question 2 Which districts experienced the greatest percentage change: Enrolment in primary schools (2008-2009) to Enrolment in primary schools (2010-2011) Average percentage change of enrollment in primary school?**

**20.**   Go to WORKING Sheet
**21.**   Select all data
**22.**   In the Main Menu go to Insert – Table – Pivot Table and Click OK in the Create PivotTable window
**23.**   Rename the new Sheet as Q2PV (Question 2 PivotTable)
**24.**   Click and drag the District/Agency field in the Rows panel
**25.**   Click and drag Enrolment in primary schools (2008-2009) and Enrolment in primary schools (2010-2011) fields in Values panel
**26.**   Change Values Fields Setting from Count to Sum in both fields, Enrolment in primary schools (2008-2009) and Enrolment in primary schools (2010-2011)

27.  It is not advisable to do additional calculation in the PivotTable. Therefore, make a new sheet and rename it Q2

28.  Go to Q2PT sheet copy all the data inside the Pivot Table, **exclude rows (blank) and GrandTotal**

29.  In Q2 sheet go to the Main Menu Home – Clipboard – Paste - Values
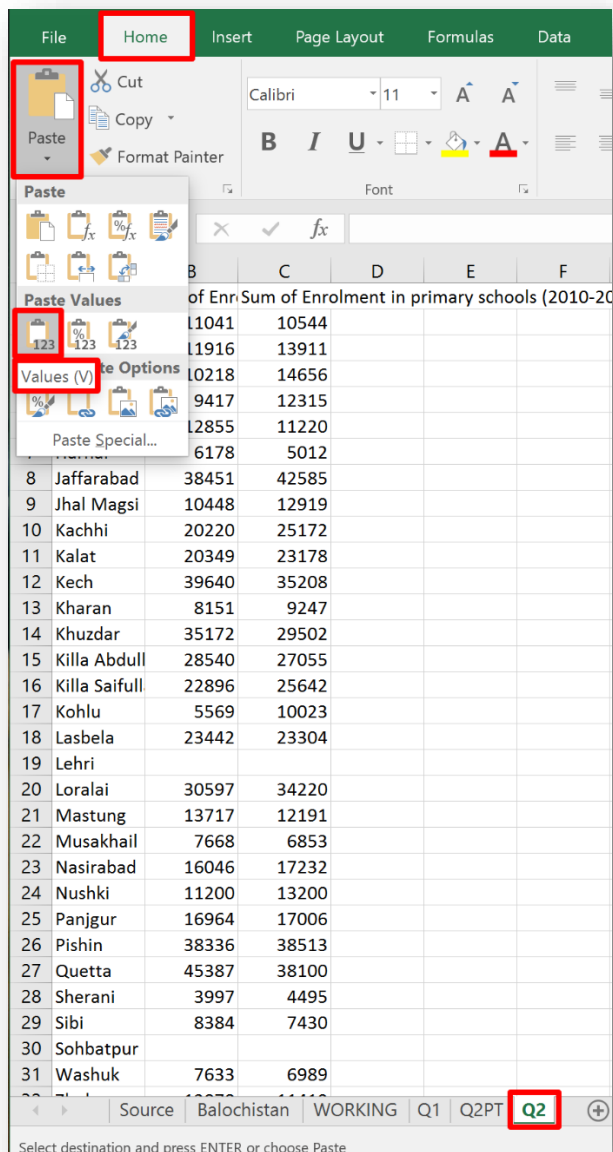
**30.**     Rename cell D1 as Percentage change 08/09 to 10/11

**31.**     To calculate the percentage change between two numbers – you need to first work out the change between the two numbers you are comparing. Then, divide the changed amount by the original number and right click, format cells and select percentage to change from a decimal to percentage. In this case you need to type **=(C2-B2)/B2**.

**32.**  Extend the formula for the rest of the districts

**33.**  Right-click in any cell and in the option panel click on format cell and in the format cell window click on Percentage.



**34.**  Now select row 1, go to the Main Menu – Home – Editing – Sort and Filter – Filter

**35.**    Go to the small arrow in D column and Sort Smallest to Largest
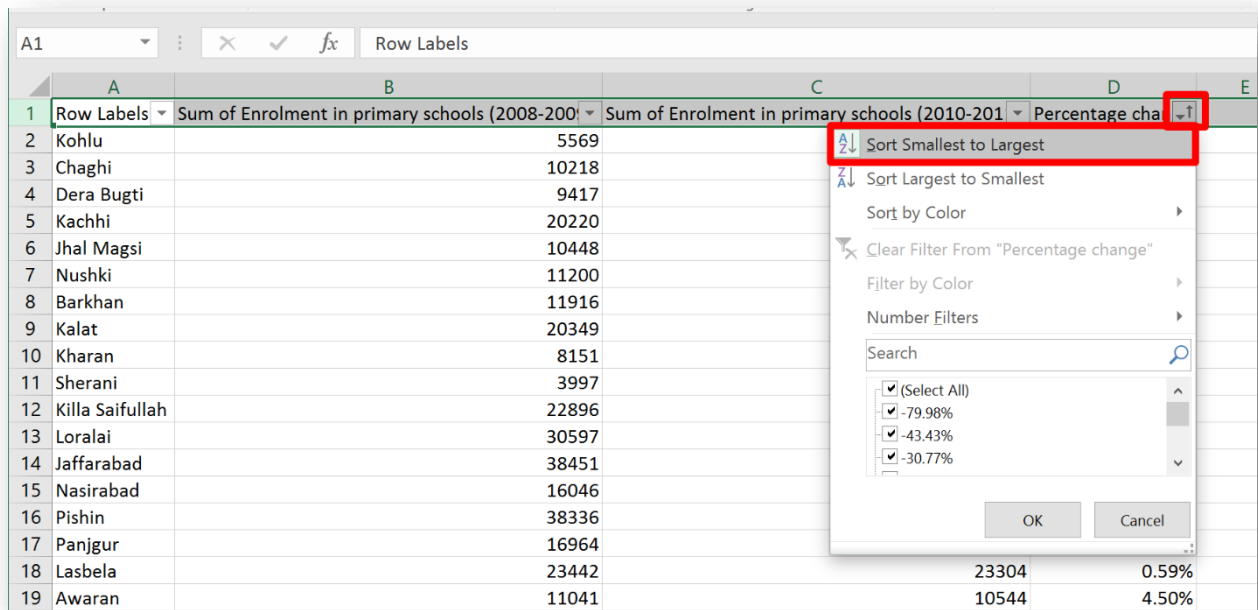


**36.**    To calculate the average percentage change of enrollment in primary school, go to cell A36 and label it average percentage change of enrollment in primary school, then in cell D36 apply the formula for average. Therefore, D36 **=AVERAGE(D2:D31).** Delete cell D32 and D33 to avoid mistakes.

**37.**    **Question 2 Which districts experienced the greatest percentage change: Enrolment in primary schools (2008-2009) to Enrolment in primary schools (2010-2011) Average percentage change of enrollment in primary school?**

| Row Labels | Sum of Enrolment in primary schools (2008-2009) | Sum of Enrolment in primary schools (2010-2011) | Percentage change |
|---|---|---|---|
| Harnai | 6178 | 5012 | -18.87% |
| Khuzdar | 35172 | 29502 | -16.12% |
| Quetta | 45387 | 38100 | -16.06% |
| Gwadar | 12855 | 11220 | -12.72% |
| Ziarat | 8282 | 7308 | -11.76% |

| | | | | |
|---|---|---|---|---|
| Sibi | 8384 | 7430 | -11.38% | |
| Zhob | 12878 | 11419 | -11.33% | |
| Kech | 39640 | 35208 | -11.18% | |
| Mastung | 13717 | 12191 | -11.12% | |
| Musakhail | 7668 | 6853 | -10.63% | |
| Washuk | 7633 | 6989 | -8.44% | |
| Killa Abdullah | 28540 | 27055 | -5.20% | |
| Awaran | 11041 | 10544 | -4.50% | |
| Lasbela | 23442 | 23304 | -0.59% | |
| Panjgur | 16964 | 17006 | 0.25% | |
| Pishin | 38336 | 38513 | 0.46% | |
| Nasirabad | 16046 | 17232 | 7.39% | |
| Jaffarabad | 38451 | 42585 | 10.75% | |
| Loralai | 30597 | 34220 | 11.84% | |
| Killa Saifullah | 22896 | 25642 | 11.99% | |
| Sherani | 3997 | 4495 | 12.46% | |
| Kharan | 8151 | 9247 | 13.45% | |
| Kalat | 20349 | 23178 | 13.90% | |
| Barkhan | 11916 | 13911 | 16.74% | |
| Nushki | 11200 | 13200 | 17.86% | |
| Jhal Magsi | 10448 | 12919 | 23.65% | |
| Kachhi | 20220 | 25172 | 24.49% | |
| Dera Bugti | 9417 | 12315 | 30.77% | |
| Chaghi | 10218 | 14656 | 43.43% | |
| Kohlu | 5569 | 10023 | 79.98% | |
| Lehri | | | | |
| Sohbatpur | | | | |
| | | | | |
| | | | | |
| | | average percentage change for all school | 5.65% | |

**Question 3 What was the percentage change in Primary schools (2008-2009) and Primary schools (2010-2011) Average percentage change of primary schools?**

**38.** Go to WORKING Sheet
**39.** Select all data
**40.** In the Main Menu go to Insert – Table – Pivot Table and Click OK in the Create PivotTable window
**41.** Rename the new Sheet as Q3PV (Question 3 PivotTable)
**42.** Click and drag District/Agency field in the Rows panel
**43.** Click and drag Primary schools (2008-2009) and Primary schools (2010-2011) fields in Values panel
**44.** Change Values Fields Setting from Count to Sum in both fields, Primary schools (2008-2009) and Primary schools (2010-2011)

**45.** Make a new sheet and rename it Q3
**46.** Go to Q3PT sheet copy all the data inside the Pivot Table, **exclude rows (blank) and Grand Total**
**47.** In Q3 sheet go to the Main Menu Home – Clipboard – Paste - Values
**48.** Rename cell D1 as Percentage change 08/09 to 10/11
**49.** To calculate the percentage, change type **=(C2-B2)/B2**.
**50.** Extend the formula for the rest of the districts
**51.** Right-click in any cell and in the option panel click on format cell and in the format cell window click on Percentage.
**52.** Now select row 1, go to the Main Menu – Home – Editing – Sort and Filter – Filter
**53.** Go to the small arrow in D column and Sort Largest to Smallest
**54.** To calculate the average percentage change of Primary schools, go to cell A36 and label it average percentage change of primary school , then in cell D36 apply the formula for average. Therefore, D36 **=AVERAGE(D2:D31).** Delete cell D32 and D33 to avoid mistakes.
**55.** **Question 3 What was the percentage change in Primary schools (2008-2009) and Sum of Primary schools (2010-2011) Average percentage change of primary school?**

| Row Labels | Sum of Primary schools (2008-2009) | Sum of Primary schools (2010-2011) | Percentage change |
|---|---|---|---|
| Kohlu | 347 | 374 | 7.78% |
| Nasirabad | 399 | 423 | 6.02% |
| Jhal Magsi | 227 | 239 | 5.29% |
| Sherani | 157 | 162 | 3.18% |
| Killa Saifullah | 521 | 534 | 2.50% |
| Killa Abdullah | 409 | 417 | 1.96% |
| Mastung | 294 | 299 | 1.70% |
| Loralai | 605 | 614 | 1.49% |
| Jaffarabad | 816 | 828 | 1.47% |
| Kachhi | 366 | 371 | 1.37% |
| Musakhail | 253 | 256 | 1.19% |
| Dera Bugti | 278 | 280 | 0.72% |
| Nushki | 151 | 152 | 0.66% |
| Ziarat | 222 | 222 | 0.00% |
| Awaran | 213 | 212 | -0.47% |
| Sibi | 207 | 206 | -0.48% |
| Kharan | 176 | 175 | -0.57% |
| Kech | 488 | 485 | -0.61% |
| Harnai | 152 | 151 | -0.66% |
| Chaghi | 196 | 194 | -1.02% |
| Gwadar | 213 | 210 | -1.41% |
| Lasbela | 489 | 482 | -1.43% |
| Panjgur | 283 | 278 | -1.77% |
| Washuk | 141 | 138 | -2.13% |
| Khuzdar | 592 | 579 | -2.20% |
| Zhob | 290 | 283 | -2.41% |
| Barkhan | 587 | 571 | -2.73% |

| | | | |
|---|---|---|---|
| Kalat | 388 | 377 | -2.84% |
| Quetta | 417 | 404 | -3.12% |
| Pishin | 791 | 752 | -4.93% |
| Lehri | | | |
| Sohbatpur | | | |
| | | | |
| | | | |
| | | average percentage change of primary school | 0.22% |

**Question 4 What was the percentage change in teaching staff from teaching staff in primary schools (2008-2009) teaching staff in primary schools (2010-2011) Average teaching staff change for all schools?**

56.    Go to WORKING Sheet
57.    Select all data
58.    In the Main Menu go to Insert – Table – Pivot Table and Click OK in the Create PivotTable window
59.    Rename the new Sheet as Q4PV (Question 4 PivotTable)
60.    Click and drag District/Agency field in the Rows panel
61.    Click and drag from teaching staff in primary schools (2008-2009) and teaching staff in primary schools (2010-2011) fields in Values panel
62.    Change Values Fields Setting from Count to Sum in both fields, from teaching staff in primary schools (2008-2009) and teaching staff in primary schools (2010-2011)
63.    Make a new sheet and rename it Q4
64.    Go to Q4PT sheet copy all the data inside the Pivot Table, **exclude rows (blank) and Grand Total**
65.    In Q4 sheet go to the Main Menu Home – Clipboard – Paste - Values
66.    Rename cell D1 as Percentage change
67.    To calculate the percentage, change type **=(C2-B2)/B2**.
68.    Extend the formula for the rest of the districts
69.    Right-click in any cell and in the option panel click on format cell and in the format cell window click on Percentage.
70.    Now select row 1, go to the Main Menu – Home – Editing – Sort and Filter – Filter
71.    Go to the small arrow in D column and Sort Largest to Smallest
72.    To calculate the average percentage change of teaching staff in primary schools, go to cell A36 and named as average percentage change of teaching staff 08/09 to 10/11, then in cell D36 apply the formula for average. Therefore, D36 **=AVERAGE(D2:D31).** Delete cell D32 and D33 to avoid mistakes.
73.    **Question 4 What was the percentage change in teaching staff from teaching staff in primary schools (2008-2009) teaching staff in primary schools (2010-2011) Average teaching staff change for all schools?**

| Row Labels | Sum of teaching staff in primary schools (2008-2009) | Sum of teaching staff in primary schools (2010-2011) | Percentage change |
|---|---|---|---|
| Washuk | 165 | 214 | 29.70% |
| Gwadar | 303 | 362 | 19.47% |
| Nushki | 296 | 353 | 19.26% |
| Sherani | 178 | 212 | 19.10% |

| | | | |
|---|---|---|---|
| Kech | 1093 | 1243 | 13.72% |
| Dera Bugti | 501 | 565 | 12.77% |
| Kohlu | 409 | 455 | 11.25% |
| Chaghi | 284 | 315 | 10.92% |
| Jhal Magsi | 376 | 413 | 9.84% |
| Mastung | 512 | 562 | 9.77% |
| Nasirabad | 601 | 656 | 9.15% |
| Panjgur | 536 | 583 | 8.77% |
| Musakhail | 329 | 357 | 8.51% |
| Lasbela | 798 | 865 | 8.40% |
| Khuzdar | 983 | 1059 | 7.73% |
| Kalat | 654 | 703 | 7.49% |
| Sibi | 485 | 521 | 7.42% |
| Kharan | 234 | 251 | 7.26% |
| Jaffarabad | 1282 | 1369 | 6.79% |
| Quetta | 1720 | 1833 | 6.57% |
| Kachhi | 700 | 745 | 6.43% |
| Zhob | 534 | 563 | 5.43% |
| Killa Saifullah | 825 | 858 | 4.00% |
| Awaran | 331 | 344 | 3.93% |
| Loralai | 922 | 958 | 3.90% |
| Killa Abdullah | 637 | 658 | 3.30% |
| Ziarat | 328 | 338 | 3.05% |
| Barkhan | 730 | 740 | 1.37% |
| Pishin | 1358 | 1347 | -0.81% |
| Harnai | 256 | 243 | -5.08% |
| Lehri | | | |
| Sohbatpur | | | |
| | | | |
| | | | |
| | | percentage change of teaching staff | 8.65% |

**Question 5 Which districts have the highest:**
**Percentage of Children of class 5 who can read letters in urdu**
**Percentage of Children of class 5 who can read words in urdu**
**Percentage of Children of class 5 who can read sentences in urdu**
**Percentage of Children of class 5 who can read story in urdu**

**74.**    Go to WORKING Sheet
**75.**    Select all data
**76.**    In the Main Menu go to Insert – Table – Pivot Table and Click OK in the Create PivotTable window
**77.**    Rename the new Sheet as Q5

**78.**    Click and drag District/Agency field in the Rows panel
**79.**    Click and drag into Values panel
- Percentage of Children of class 5 who can read letters in urdu
- Percentage of Children of class 5 who can read words in urdu
- Percentage of Children of class 5 who can read sentences in urdu
- Percentage of Children of class 5 who can read story in urdu

**80.**    Change Values Fields Setting from Count to Sum in all fields added
**81.**    Make a new sheet and rename it Q5
**82.**    Go to Q5PT sheet copy all the data inside the Pivot Table, **exclude rows (blank) and Grand Total**
**83.**    In Q5 sheet go to the Main Menu Home – Clipboard – Paste – Values
**84.**    Using Wrap Text to arrange Row 1 in the way to visualize better the data. By select Row 1 go to the Main Menu – Home – Alignment – Wrap Text



**85.**    Add a filter and filter each one at a time:
- read letters first
- read words second
- read sentences third
- read story fourth

**86.    Question 5 Which districts have the highest:**
**Percentage of Children of class 5 who can read letters in urdu**
**Percentage of Children of class 5 who can read words in urdu**
**Percentage of Children of class 5 who can read sentences in urdu**
**Percentage of Children of class 5 who can read story in urdu**

| Row Labels | Sum of Percentage of Children of class 5 who can read letters in urdu | Sum of Percentage of Children of class 5 who can read words in urdu | Sum of Percentage of Children of class 5 who can read sentences in urdu | Sum of Percentage of Children of class 5 who can read story in urdu |
|---|---|---|---|---|
| Musakhail | 57 | 17.4 | 8.3 | 6.6 |
| Mastung | 26.7 | 43.8 | 15.2 | 1.9 |
| Chaghi | 20.6 | 17.6 | 41.2 | 18.2 |
| Sibi | 17.8 | 42.5 | 24.7 | 13.7 |
| Loralai | 9.1 | 30.7 | 40.9 | 13.6 |
| Killa Abdullah | 6.5 | 51.1 | 18.5 | 21.7 |
| Harnai | 6.1 | 18.2 | 42.4 | 27.3 |
| Kachhi | 5.8 | 11.5 | 13.5 | 65.4 |
| Kech | 2.5 | 51.6 | 29.8 | 14.3 |
| Panjgur | 2.4 | 12.1 | 57.6 | 27.3 |
| Khuzdar | 2.3 | 4.7 | 37.2 | 55.8 |
| Kalat | 1.8 | 12.5 | 37.5 | 48.2 |
| Pishin | 1.7 | 16.4 | 20.5 | 59.7 |
| Kharan | 1.1 | 5.6 | 11.1 | 80.6 |
| Barkhan | 0.7 | 12.1 | 53.6 | 32.9 |
| Nasirabad | 0.7 | 1.4 | 9.9 | 87.3 |
| Quetta | 0.6 | 4.1 | 18 | 77.3 |
| Dera Bugti | 0 | 60 | 20 | 20 |
| Nushki | 0 | 45.9 | 48.6 | 5.4 |
| Gwadar | 0 | 31.9 | 2.7 | 65.4 |
| Lasbela | 0 | 14.3 | 67.9 | 17.9 |
| Sherani | 0 | 10.3 | 24.1 | 65.5 |
| Kohlu | 0 | 7 | 30.2 | 60.5 |
| Zhob | 0 | 3 | 57.6 | 39.4 |
| Jhal Magsi | 0 | 1.7 | 6.7 | 91.7 |
| Jaffarabad | 0 | 0.7 | 2 | 97.4 |
| Ziarat | 0 | 0 | 32.3 | 67.7 |
| Killa Saifullah | 0 | 0 | 1.4 | 97.1 |

**Additional Exercises**
**1.**    Question 6 Which districts have the highest:
- Percentage of children of class 5 who can recognize number from 1-9
- Percentage of children of class 5 who can recognize number from 10-99

- Percentage of children of class 5 who can subtract two digits
- Percentage of children of class 5 who can divide two digits


2. The dataset use in the exercise also can also answer different question related with the Hypothesis and the quality of the education in primary school. As example:
   - The number of students for each teacher in primary school
   - The number of teachers per school in primary school
   - The number of students per school
   - The last three calculations can be done by year to show if the quality has been improving or not.