Integrated project: Access to drinking water (Part 2 - Transforming the data) [MCQ] (Version : 0)

TEST

Correct Answer

Answered in 5.3166666666667 Minutes

Question 1/10

True or false? The dataset provided represents the entire year range from 2000 to 2020 for at least some countries.

- [ ] True
- [x] False

Explanation:

Although the CSV is called “Estimated water use of 2000 - 2020.csv”, considering the minimum and maximum (and average) value of the year column, the dataset only represents the top quartile of the range, i.e. the years 2015 to 2020.

Question 2/10

What is the average year difference across all countries for the dataset?

- [ ] 4.34 years
- [ ] 4.36 years
- [ ] 4.76 years
- [x] 4.80 years

Explanation:

The correct value is found only if the IF statement is correct, and after the two duplicate rows for Zimbabwe are removed. The average of 4.76 is found if the IF statement is correct but the duplicates have not been removed. 4.34 is found if the duplicates are not removed and the IF statement is not used (i.e. subsequent years are simply subtracted from each other without checking for equality between names). Similarly, 4.36 is found if the duplicates are removed but the IF statement is incorrect.

Question 3/10

Which of the following observations best represents the distribution of the year column?

- [ ] The distribution is positively skewed since the peak is to the left of the number line.
- [x] The distribution is neither normal, negatively, nor positively skewed since there are two distinct peaks.
- [ ] Nothing can be said about the distribution because we have too few data points.
- [ ] The distribution is negatively skewed since the peak is to the left of the number line.

Explanation:

Without changing the x-axis minimum and maximum values and the bucket size, nothing can be said about the distribution because all values will be collectively shown within one column. Without ensuring that the maximum is set to the maximum year plus one, it will seem like the peak is to the left of the number line because the second peak is in 2020.

Question 4/10

What is the average Annual Rates of Change (ARC) of access to basic water services for rural populations (ARC_r) across all countries?

- [ ] 0.3503%
- [ ] 0.0242%
- [x] 0.48%
- [ ] 48.45%

Explanation:

If the relative reference is kept in the numerator of the ARC formula, the calculation would be based on wat_lim_n and not wat_bas_r, and the answer would then incorrectly be 0.0242. If the error values are replaced with zeros rather than ”null”, the answer would incorrectly be 0.3503. Since we are already taking the difference between values in percentage points (and dividing by year difference) to calculate ARC, the ARC value is already in percentage points (i.e. it shouldn’t be multiplied by 100%), and therefore, 48.45% is incorrect.

Question 5/10

How many countries’ national populations had a 0% Annual Rates of Change, excluding countries that have 100% access, across the time period?

- [x] 16
- [ ] 62
- [ ] 14
- [ ] 21

Explanation:

The answer 16 is determined through COUNTIFS. The first criteria range is ARC_n with a criterion of

“=0”, and the second range is ARC_n_full with a criterion “<>full access”. In other words, only count when ARC_n = 0 AND NOT full access. If the rounding is done incorrectly, for example, rounding down with two decimal places, the answer would incorrectly be 21. The number of countries where national population access is full (i.e. wat_bas_n == 100%) is equal to 62, not the number of countries with a 0% ARC.

Question 6/10

Considering that a negative Annual Rates of Change (ARC) indicates a decrease in access to water from one year to the next, and a zero ARC indicates no change, which of the following statements is most true based on the data?

- [x] Although access to basic water services on a national level increased for more countries, more countries had a decrease in access in urban than rural areas.
- [ ] Although access to basic water services on a national level remained unchanged for more countries than in rural and urban areas, rural areas had the greatest decrease in access.
- [ ] Although access to basic water services increased for more countries in rural areas than on a national level, more countries had full access to basic water services in urban areas.
- [ ] Most countries had a similar change in access to basic water services across all types of population areas.

Explanation:

This question requires an interpretation of the number of countries that saw a decrease, increase, or unchanged ARC value. We can say that access on a national level increased more since 135 countries had ARC_n > 0, while in rural areas ARC_r > 0 is 116, and in urban areas ARC_u > 0 is 93, although the average of ARC_r is the highest.

This means that although fewer countries saw an increase in rural areas than on a national level, the percentage change per country was on average higher in rural areas. Also, while ARC_u < 0 is 26 countries, ARC_r is 17 countries, which means fewer countries saw a decrease in rural areas.

Question 7/10

Which two countries had the highest *absolute difference* between urban and rural Annual Rates of Change?

- [ ] South Sudan and Burkina Faso
- [x] South Sudan and Morocco
- [ ] Afghanistan and Morocco
- [ ] Morocco and Iraq

Explanation:

The question is about the absolute difference in ARC, so subtracting ARC_r from ARC_u, or the other way around, will not change the answer. We've only asked them in the instructions to calculate the difference (not the absolute difference) to see if they can interpret differences, and then consider percentage points as the unit.

Question 8/10

Which of the following statements is most true about the distribution of the difference in ARC values between rural and urban areas?

- [ ] The Annual Rates of Change were similar across urban and rural areas for most countries since the peak is close to the middle of the number line.
- [ ] More countries had higher Annual Rates of Change in urban areas than in rural areas

since a greater number of difference values falls to the one side of the number line.

We cannot say anything about the distribution because some of the histogram bins are empty.

- [x] More countries had higher Annual Rates of Change in rural areas than in urban areas since a greater number of difference values falls to the one side of the number line.

Explanation:

If we calculate ARC_diff as ARC_r - ARC_u, we see that more of the difference values fall to the right of the number line, i.e. for more calculated differences, the difference was positive indicating that the rural ARC values were greater than the urban ARC values. If we were to calculate the difference as ARC_u - ARC_r, we would find more negative values which still indicates that more rural ARC values were greater.

Question 9/10

On average, which region saw the greatest improvement in access to basic water services on a national level (considering the Annual Rates of Change) over the dataset time period?

- [ ] South Asia
- [x] Sub-Saharan Africa
- [ ] North America
- [ ] Middle East & North Africa

Explanation:

Using a pivot table, we can calculate the average of ARC_n to determine the highest value, which is Sub-Saharan Africa for AVERAGE of ARC_n.

Question 10/10

Based on the visualisation investigating the relationship between the rural and national Annual Rates of Change, national population size, and region, which of the following statements are true?

a. Only countries in the Sub-Saharan Africa region observed a decrease in basic water access on a national level and in rural areas.
b. The average population size per country in Sub-Saharan Africa is smaller than that of most other regions.
c. Countries in the Sub-Saharan Africa region observed a greater spread in rural and national ARC values than other regions.
d. Countries with larger populations generally observed national ARC values between 0% and 1%.

- [ ] Only a, b, and c
- [x] Only b, c, and d
- [ ] All of the above
- [ ] Only b

Explanation:

Option a is false; there are at least three countries in the East Asia & Pacific region and two in the Europe & Central Asia region also observed a decrease in both ARC_n and ARC_r. Option b is true; although the total population of the Sub-Saharan Africa region is not the smallest, it includes more countries than most regions (except Europe & Central Asia), which means that the population is more distributed across the different countries. A similar observation can be made about the Europe & Central Asia region. Option c is true; countries in the Sub-Saharan Africa region observed ARC_n and ARC_r values between approximately -2 and 3, while East Asia & Pacific region lies approximately between -1 and 2.5, and South Asia between 0 and 3. Option d is also true; countries with large populations (mostly in the South Asia and East Asia & Pacific regions) generally have national ARC values between 0 and 1%.