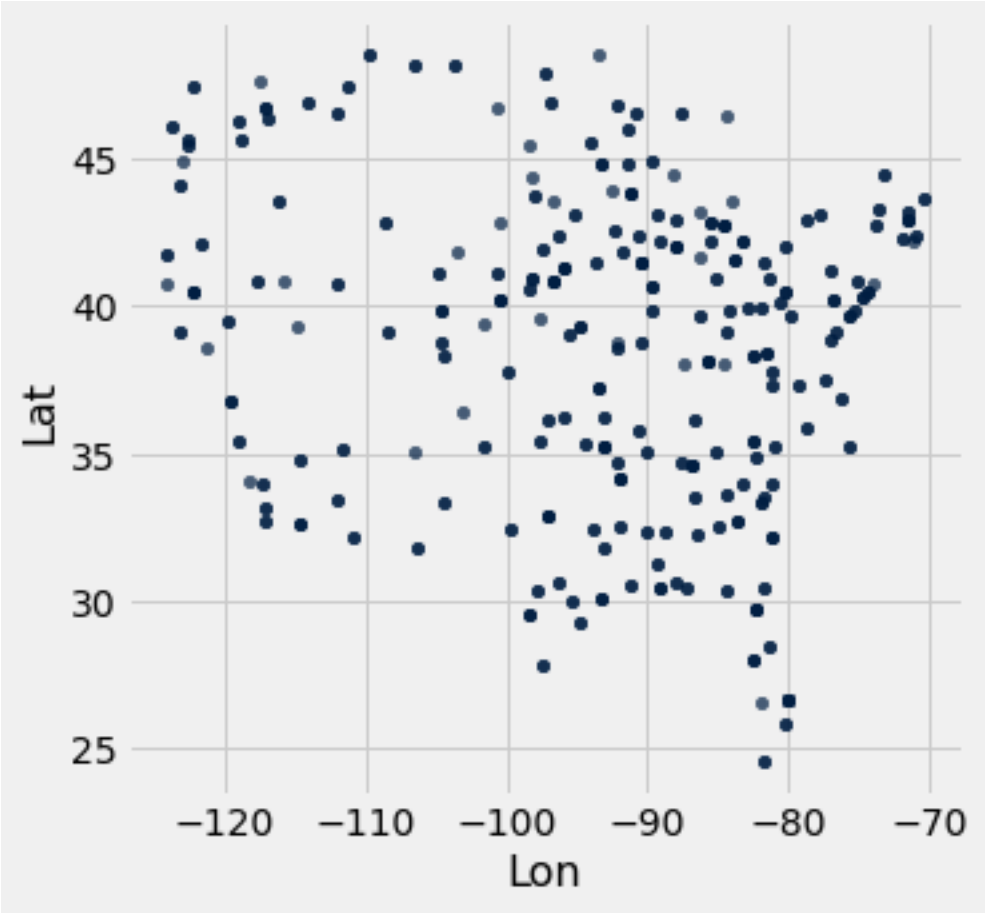
Question 1.1.1: In the cell below, produce a scatter plot that plots the latitude and longitude of every city in the `cities` table so that the result places northern cities at the top and western cities at the left.

Note: It's okay to plot the same point multiple times!

```
In [3]: cities.scatter("Lon", "Lat") # SOLUTION
```



Question 1.1.2 Does it appear that these city locations are sampled uniformly at random from all the locations in the U.S.? Why or why not?

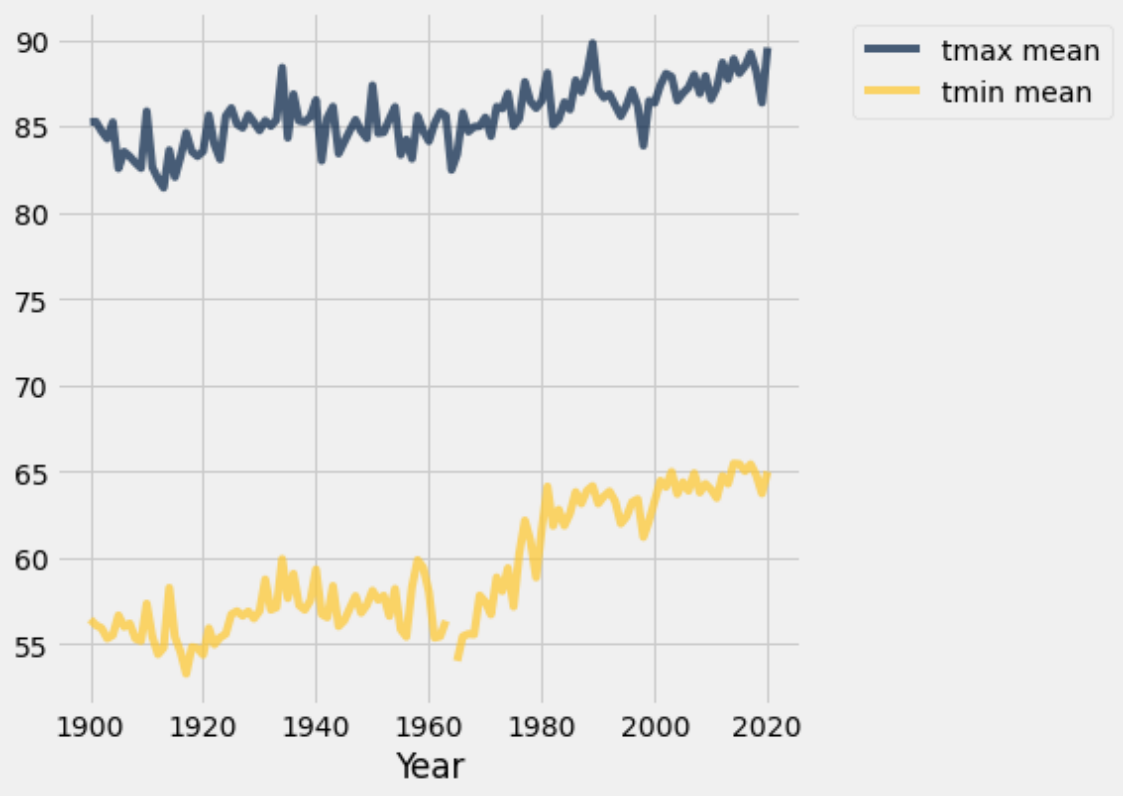
Type your answer here, replacing this text.

SOLUTION: No. The cities are very concentrated around the coasts; a uniform random sample would include more cities in the central regions.

Question 1.2.3: Using the `phoenix` table, create an overlaid line plot of the **average maximum temperature** and **average minimum temperature** for each year between 1900 and 2020 (inclusive).

Hint: To draw a line plot with more than one line, call `plot` on the column label of the x-axis values and all other columns will be treated as y-axis values.

```
In [38]: phoenix.select('Year', 'tmax', 'tmin').group("Year", np.mean).where('Year', are.between_or_equa
```



Question 1.2.4: Although still hotly debated (pun intended), many climate scientists agree that the effects of climate change began to surface in the early 1960s as a result of elevated levels of greenhouse gas emissions. How does the graph you produced in Question 1.2.3 support the claim that modern-day global warming began in the early 1960s?

Type your answer here, replacing this text.

SOLUTION: After about 1963, the minimum temperatures increase rapidly and the maximum temperatures increase slowly. Before then, they stayed within a consistent range.

Question 1.2.7 The lower bound of the `feb_present_ci` 99% confidence interval is below the observed past February average maximum temperature of 68.8485 (from the `monthly_increases` table). What conclusion can you draw about the effect of climate change on February maximum temperatures in Phoenix from this information? Use a 1% p-value cutoff.

Note: If you're stuck on this question, re-reading the paragraphs under the *February* heading (particularly the first few) may be helpful.

Type your answer here, replacing this text.

SOLUTION: The confidence interval for the present average contains the past average, so we cannot reject the null hypothesis that the present average is in fact the past average. Under our assumptions, it's consistent with the data that climate change has not affected the average maximum February temperature at all.

Question 1.2.9. Summarize your findings. After comparing the past average to the 99% confidence interval's lower bound for each month, what conclusions can we make about the monthly average maximum temperature in historical (1900-1960) vs. modern (2019-2021) times in the twelve months? In other words, what null hypothesis should you consider, and for which months would you reject or fail to reject the null hypothesis? Use a 1% p-value cutoff.

Hint: Do you notice any seasonal patterns?

Type your answer here, replacing this text.

SOLUTION: For most months, the modern average maximum temperature for the month is above the historical average maximum temperature, indicating the climate has changed. But for some winter months—Feb, March, Dec—we fail to reject the null, as we do not have evidence that the average maximum temperature has changed.

Question 2.2. Define null and alternative hypotheses for an A/B test that investigates whether drought years are drier (have less precipitation) than other years.

Note: Please format your answer using the following structure.

- *Null hypothesis: …*
- *Alternative hypothesis: …*

Type your answer here, replacing this text.

SOLUTION:

- *Null hypothesis:* The drought year total precipitation amounts are sampled from the same distribution as other year total precipitation amounts.
- *Alternative hypothesis:* The drought year total precipitation amounts are sampled from a drier distribution than the other year total precipitation amounts.

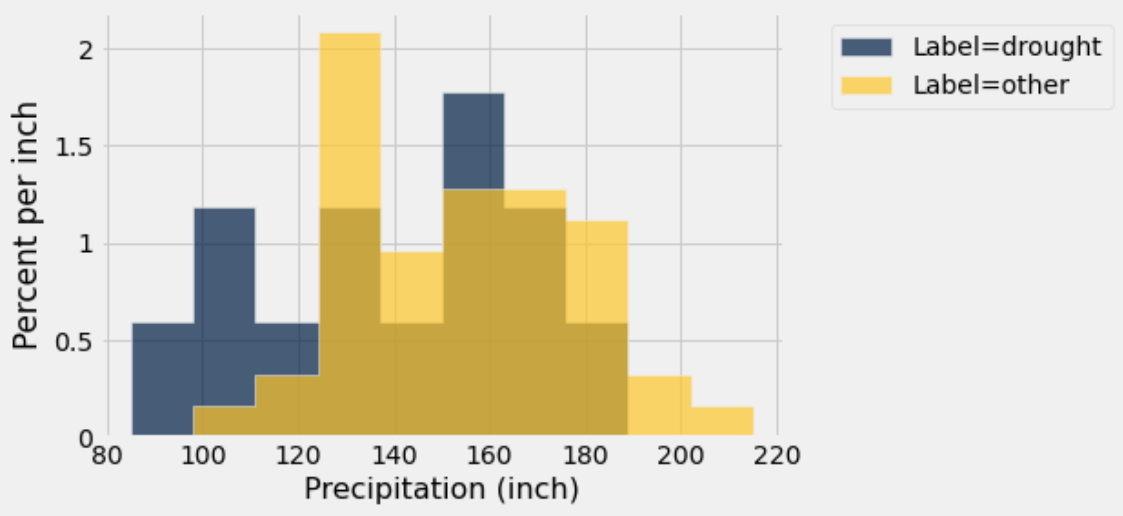
Question 2.3. First, define the table `drought`. It should contain one row per year and the following two columns: - `"Label"`: Denotes if a year is part of a `"drought"` year or an `"other"` year - `"Precipitation"`: The sum of the total precipitation in 13 Southwest cities that year

Then, construct an overlaid histogram of two observed distributions: the total precipitation in drought years and the total precipitation in other years.

Note: Use the provided `bins` when creating your histogram, and do not re-assign the `southwest` table. Feel free to use as many lines as you need!

Hint: The optional `group` argument in a certain function might be helpful!

```
In [63]: bins = np.arange(85, 215+1, 13)
         drought = totals.with_column('Label', totals.apply(drought_label, 'Year')).select('Label', 'Pr
         drought.hist("Precipitation", group="Label", bins=bins, unit='inch') # SOLUTION
```



Question 2.5. Fellow climate scientists Olivia and Will point out that there are more **other** years than **drought** years, and so measuring the difference between total precipitation will always favor the **other** years. They conclude that all of the options above involving **total** precipitation are invalid test statistic choices. Do you agree with them? Why or why not?

Type your answer here, replacing this text.

SOLUTION: No. Permutation tests apply equally well to datasets with imbalanced classes, and an averaging test statistic is not required. A permutation test always ensures that there are the same number of each class during simulation, making the observed and simulated test statistics comparable.

Question 2.10. State a conclusion from this test using a p-value cutoff of 5%. What have you learned about the EPA's statement on drought?

Type your answer here, replacing this text.

SOLUTION: Because the p-value is less than the 5% p-value cutoff, our results are statistically significant and we reject the null hypothesis. We can conclude that the drought years were drier than the other years in these 13 cities.

Question 2.11. Does your conclusion from Question 2.10 apply to the entire Southwest region of the U.S.? Why or why not?

Type your answer here, replacing this text.

SOLUTION: We cannot generalize to the whole Southwestern U.S. because A/B tests only provide conclusions about the observed subjects, not the population from which they were drawn. The cities we measured may not be representative of the whole region, which includes lots of mountains and deserts that aren't particularly suitable to city construction.