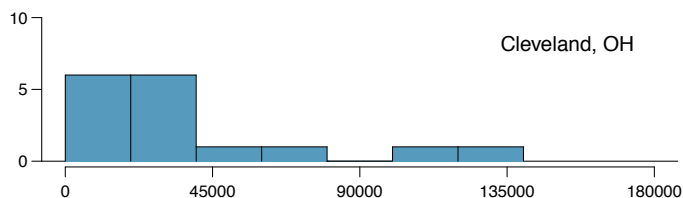
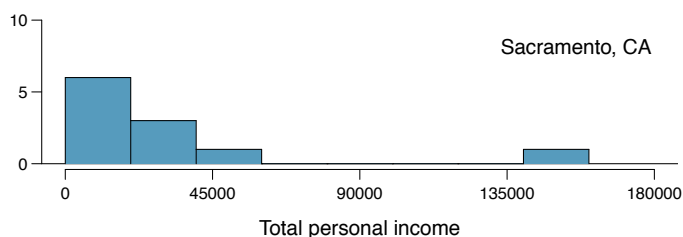


5.6.3 Difference of two means

5.25 Cleveland vs. Sacramento. Average income varies from one region of the country to another, and it often reflects both lifestyles and regional living expenses. Suppose a new graduate is considering a job in two locations, Cleveland, OH and Sacramento, CA, and he wants to see whether the average income in one of these cities is higher than the other. He would like to conduct a hypothesis test based on two small samples from the 2000 Census, but he first must consider whether the conditions are met to implement the test. Below are histograms for each city. Should he move forward with the hypothesis test? Explain your reasoning.

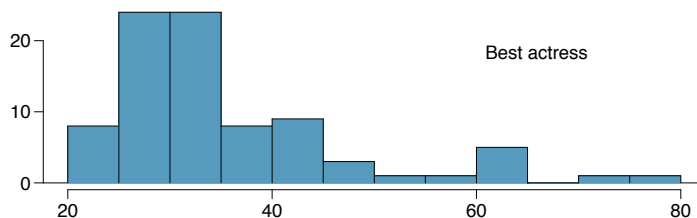


Cleveland, OH	
Mean	\$ 35,749
SD	\$ 39,421
n	21

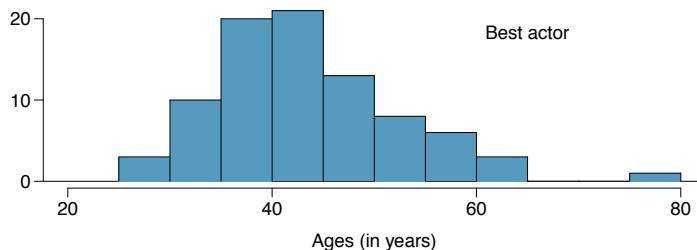


Sacramento, CA	
Mean	\$ 35,500
SD	\$ 41,512
n	17

5.26 Oscar winners. The first Oscar awards for best actor and best actress were given out in 1929. The histograms below show the age distribution for all of the best actor and best actress winners from 1929 to 2012. Summary statistics for these distributions are also provided. Is a hypothesis test appropriate for evaluating whether the difference in the average ages of best actors and actresses might be due to chance? Explain your reasoning.³⁸



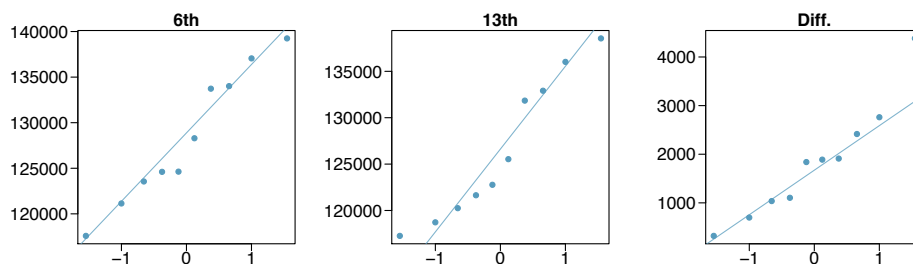
Best Actress	
Mean	35.6
SD	11.3
n	84



Best Actor	
Mean	44.7
SD	8.9
n	84

³⁸Oscar winners from 1929 – 2012, data up to 2009 from the *Journal of Statistics Education* data archive and more current data from wikipedia.org.

5.27 Friday the 13th, Part I. In the early 1990's, researchers in the UK collected data on traffic flow, number of shoppers, and traffic accident related emergency room admissions on Friday the 13th and the previous Friday, Friday the 6th. The histograms below show the distribution of number of cars passing by a specific intersection on Friday the 6th and Friday the 13th for many such date pairs. Also given are some sample statistics, where the difference is the number of cars on the 6th minus the number of cars on the 13th.³⁹



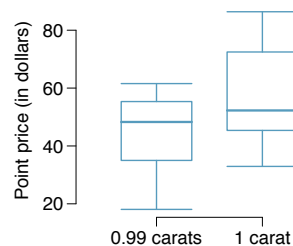
	6 th	13 th	Diff.
\bar{x}	128,385	126,550	1,835
s	7,259	7,664	1,176
n	10	10	10

- Are there any underlying structures in these data that should be considered in an analysis? Explain.
- What are the hypotheses for evaluating whether the number of people out on Friday the 6th is different than the number out on Friday the 13th?
- Check conditions to carry out the hypothesis test from part (b).
- Calculate the test statistic and the p-value.
- What is the conclusion of the hypothesis test?
- Interpret the p-value in this context.
- What type of error might have been made in the conclusion of your test? Explain.

5.28 Diamonds, Part I. Prices of diamonds are determined by what is known as the 4 Cs: cut, clarity, color, and carat weight. The prices of diamonds go up as the carat weight increases, but the increase is not smooth. For example, the difference between the size of a 0.99 carat diamond and a 1 carat diamond is undetectable to the naked human eye, but the price of a 1 carat diamond tends to be much higher than the price of a 0.99 diamond. In this question we use two random samples of diamonds, 0.99 carats and 1 carat, each sample of size 23, and compare the average prices of the diamonds. In order to be able to compare equivalent units, we first divide the price for each diamond by 100 times its weight in carats. That is, for a 0.99 carat diamond, we divide the price by 99. For a 1 carat diamond, we divide the price by 100. The distributions and some sample statistics are shown below.⁴⁰

Conduct a hypothesis test to evaluate if there is a difference between the average standardized prices of 0.99 and 1 carat diamonds. Make sure to state your hypotheses clearly, check relevant conditions, and interpret your results in context of the data.

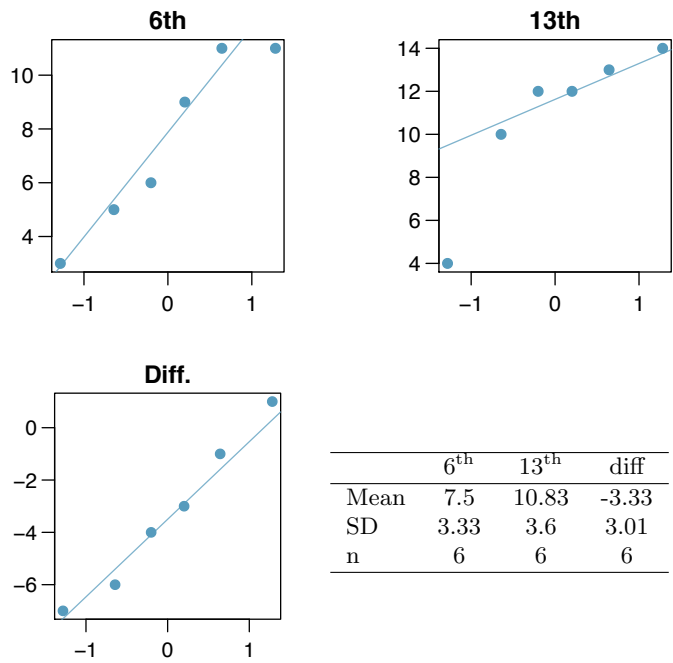
	0.99 carats	1 carat
Mean	\$ 44.51	\$ 56.81
SD	\$ 13.32	\$ 16.13
n	23	23



³⁹T.J. Scanlon et al. "Is Friday the 13th Bad For Your Health?" In: *BMJ* 307 (1993), pp. 1584–1586.

⁴⁰H. Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.

5.29 Friday the 13th, Part II. The Friday the 13th study reported in Exercise 5.27 also provides data on traffic accident related emergency room admissions. The distributions of these counts from Friday the 6th and Friday the 13th are shown below for six such paired dates along with summary statistics. You may assume that conditions for inference are met.



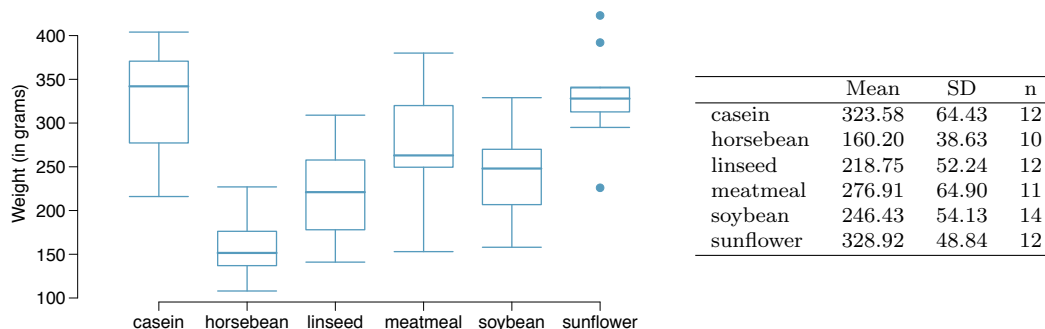
- (a) Conduct a hypothesis test to evaluate if there is a difference between the average numbers of traffic accident related emergency room admissions between Friday the 6th and Friday the 13th.
- (b) Calculate a 95% confidence interval for the difference between the average numbers of traffic accident related emergency room admissions between Friday the 6th and Friday the 13th.
- (c) The conclusion of the original study states, “Friday 13th is unlucky for some. The risk of hospital admission as a result of a transport accident may be increased by as much as 52%. Staying at home is recommended.” Do you agree with this statement? Explain your reasoning.

5.30 Diamonds, Part II. In Exercise 5.28, we discussed diamond prices (standardized by weight) for diamonds with weights 0.99 carats and 1 carat. See the table for summary statistics, and then construct a 95% confidence interval for the average difference between the standardized prices of 0.99 and 1 carat diamonds. You may assume the conditions for inference are met.

	0.99 carats	1 carat
Mean	\$ 44.51	\$ 56.81
SD	\$ 13.32	\$ 16.13
n	23	23

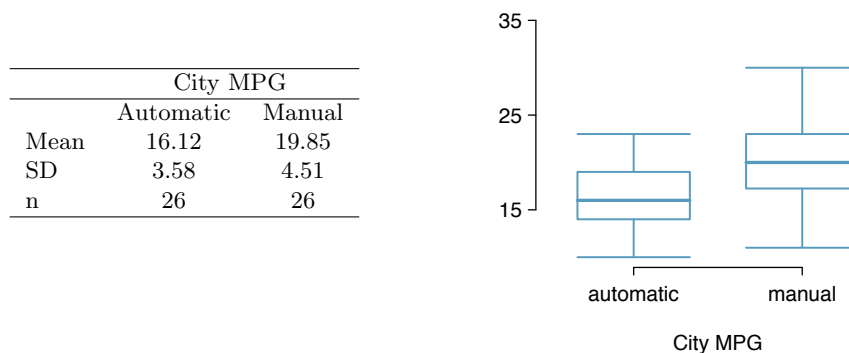


5.31 Chicken diet and weight, Part I. Chicken farming is a multi-billion dollar industry, and any methods that increase the growth rate of young chicks can reduce consumer costs while increasing company profits, possibly by millions of dollars. An experiment was conducted to measure and compare the effectiveness of various feed supplements on the growth rate of chickens. Newly hatched chicks were randomly allocated into six groups, and each group was given a different feed supplement. Below are some summary statistics from this data set along with box plots showing the distribution of weights by feed type.⁴¹



- Describe the distributions of weights of chickens that were fed linseed and horsebean.
- Do these data provide strong evidence that the average weights of chickens that were fed linseed and horsebean are different? Use a 5% significance level.
- What type of error might we have committed? Explain.
- Would your conclusion change if we used $\alpha = 0.01$?

5.32 Fuel efficiency of manual and automatic cars, Part I. Each year the US Environmental Protection Agency (EPA) releases fuel economy data on cars manufactured in that year. Below are summary statistics on fuel efficiency (in miles/gallon) from random samples of cars with manual and automatic transmissions manufactured in 2012. Do these data provide strong evidence of a difference between the average fuel efficiency of cars with manual and automatic transmissions in terms of their average city mileage? Assume that conditions for inference are satisfied.⁴²



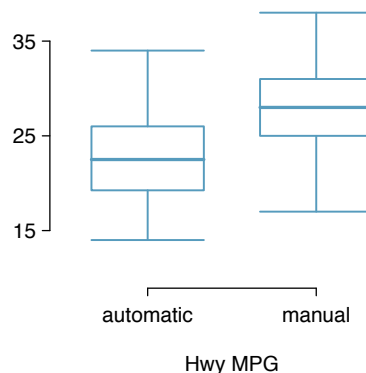
5.33 Chicken diet and weight, Part II. Casein is a common weight gain supplement for humans. Does it have an effect on chickens? Using data provided in Exercise 5.31, test the hypothesis that the average weight of chickens that were fed casein is different than the average weight of chickens that were fed soybean. If your hypothesis test yields a statistically significant result, discuss whether or not the higher average weight of chickens can be attributed to the casein diet. Assume that conditions for inference are satisfied.

⁴¹Chicken Weights by Feed Type, from the `datasets` package in R..

⁴²U.S. Department of Energy, Fuel Economy Data, 2012 Datafile.

5.34 Fuel efficiency of manual and automatic cars, Part II. The table provides summary statistics on highway fuel economy of cars manufactured in 2012 (from Exercise 5.32). Use these statistics to calculate a 98% confidence interval for the difference between average highway mileage of manual and automatic cars, and interpret this interval in the context of the data.⁴³

	Hwy MPG	
	Automatic	Manual
Mean	22.92	27.88
SD	5.29	5.01
n	26	26



5.35 Gaming and distracted eating, Part I. A group of researchers are interested in the possible effects of distracting stimuli during eating, such as an increase or decrease in the amount of food consumption. To test this hypothesis, they monitored food intake for a group of 44 patients who were randomized into two equal groups. The treatment group ate lunch while playing solitaire, and the control group ate lunch without any added distractions. Patients in the treatment group ate 52.1 grams of biscuits, with a standard deviation of 45.1 grams, and patients in the control group ate 27.1 grams of biscuits, with a standard deviation of 26.4 grams. Do these data provide convincing evidence that the average food intake (measured in amount of biscuits consumed) is different for the patients in the treatment group? Assume that conditions for inference are satisfied.⁴⁴

5.36 Gaming and distracted eating, Part II. The researchers from Exercise 5.35 also investigated the effects of being distracted by a game on how much people eat. The 22 patients in the treatment group who ate their lunch while playing solitaire were asked to do a serial-order recall of the food lunch items they ate. The average number of items recalled by the patients in this group was 4.9, with a standard deviation of 1.8. The average number of items recalled by the patients in the control group (no distraction) was 6.1, with a standard deviation of 1.8. Do these data provide strong evidence that the average number of food items recalled by the patients in the treatment and control groups are different?

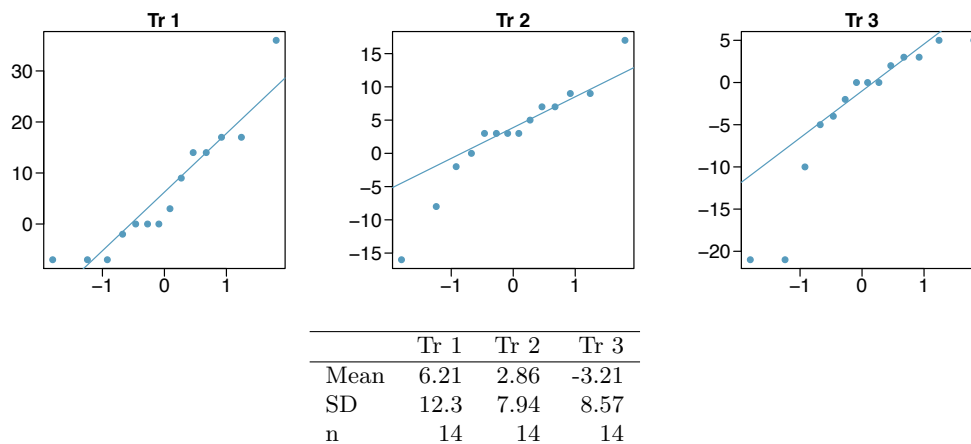
⁴³U.S. Department of Energy, Fuel Economy Data, 2012 Datafile.

⁴⁴R.E. Oldham-Cooper et al. "Playing a computer game during lunch affects fullness, memory for lunch, and later snack intake". In: *The American Journal of Clinical Nutrition* 93.2 (2011), p. 308.

5.37 Prison isolation experiment, Part I. Subjects from Central Prison in Raleigh, NC, volunteered for an experiment involving an “isolation” experience. The goal of the experiment was to find a treatment that reduces subjects’ psychopathic deviant T scores. This score measures a person’s need for control or their rebellion against control, and it is part of a commonly used mental health test called the Minnesota Multiphasic Personality Inventory (MMPI) test. The experiment had three treatment groups:

- (1) Four hours of sensory restriction plus a 15 minute “therapeutic” tape advising that professional help is available.
- (2) Four hours of sensory restriction plus a 15 minute “emotionally neutral” tape on training hunting dogs.
- (3) Four hours of sensory restriction but no taped message.

Forty-two subjects were randomly assigned to these treatment groups, and an MMPI test was administered before and after the treatment. Distributions of the differences between pre and post treatment scores (pre - post) are shown below, along with some sample statistics. Use this information to independently test the effectiveness of each treatment. Make sure to clearly state your hypotheses, check conditions, and interpret results in the context of the data.⁴⁵



5.38 True / False: comparing means. Determine if the following statements are true or false, and explain your reasoning for statements you identify as false.

- (a) When comparing means of two samples where $n_1 = 20$ and $n_2 = 40$, we can use the normal model for the difference in means since $n_2 \geq 30$.
- (b) As the degrees of freedom increases, the t -distribution approaches normality.
- (c) We use a pooled standard error for calculating the standard error of the difference between means when sample sizes of groups are equal to each other.

⁴⁵Prison isolation experiment.