GROUND RULES:

- **Print** your full name clearly at the top of this page. Use the name that appears on university records.
- This is a closed-book and closed-notes exam. You can not use external notes of any kind.
- You may use a calculator. You may not use your phone as a calculator.
- This exam contains three parts:
  - Part 1. Multiple Choice. 25 questions, 2 points each (**50 points** total)
  - Part 2. Short Answer. 1 question, 10 points each **(10 points** total)
  - Part 3. Extra Credit. 1 question (5 points total).

  This exam is worth **60 points** (but it is possible to get up to 65 points).
- Any discussion or inappropriate communication between you and another examinee, as well as the appearance of any unnecessary material, is not allowed. All violations will be reported to the Student Conduct and Academic Integrity Office immediately.
- You have **50 minutes** to complete this exam.

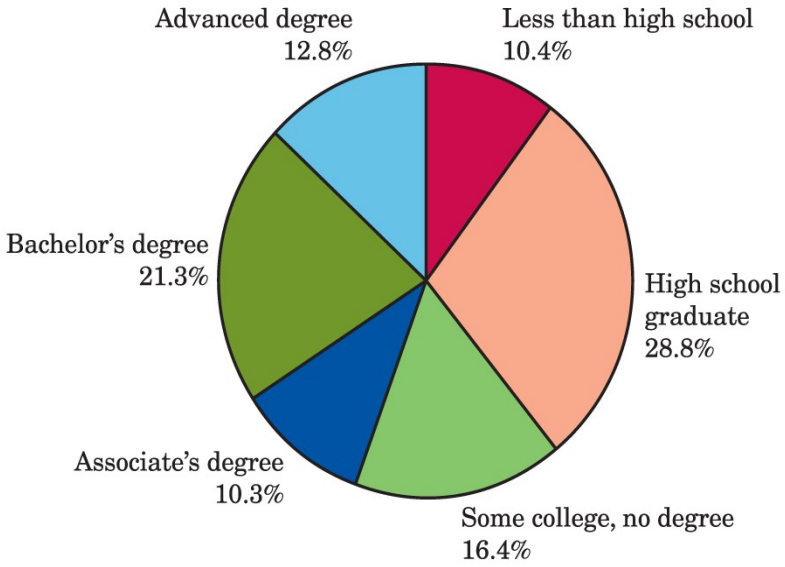
HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

> *I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.*

PART 1: MULTIPLE CHOICE. Circle the best answer. Make sure your answer is clearly marked. Ambiguous responses will be marked wrong.

1. Moore and Notz (2020) show the following pie chart to display the distribution of the level of education of people aged 25 years and older in the United States in 2017:



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020
W. H. Freeman and Company

A pie chart is used because level of education is a _____ variable.

(a) categorical
(b) quantitative

2. In class, we talked about the **Tuskegee syphilis study** which started in 1932 and involved 600 black men: 399 with syphilis and 201 who did not have the disease.

What one good thing came from study?

(a) The study investigators used the results to create new approaches to perform statistical analyses.
(b) A cure for syphilis was found. The study's two researchers won the Nobel Prize for Medicine.
(c) The US government created new laws mandating strict guidelines for human subjects research.
(d) This study helped to shape how sex education is taught in today's high schools.

3. What does the **distribution** of a variable tell us?

(a) whether the variable is measured with bias and random error
(b) what values the variable can have and how often it has these values
(c) whether the variable has predictive validity
(d) what type of study was used to measure the variable

4. Last semester, a student concerned about his performance in my class made a 40 on the first exam. He made a 50 on the second exam. What is the **percentage change** in his exam scores?

(a) 20%
(b) 10%
(c) 90%
(d) 25%

5. When we did the FATHERLY FRIENDLY FARMS example in class, one student counted the number of F's to be 14, 16, and 23. The true number of F's is 48. This student's measurements suffered from

(a) bias and random error.
(b) random error only.
(c) bias only.

6. On October 8, 2024, the American Automobile Association (AAA) reported the five states where gas is the most expensive in the United States. Here are the states and the average gas prices (dollars/gallon) for each state:

State	Average price
California	\$4.67
Hawaii	\$4.61
Washington	\$4.02
Nevada	\$3.90
Alaska	\$3.66

Which graph should we use to display this information?

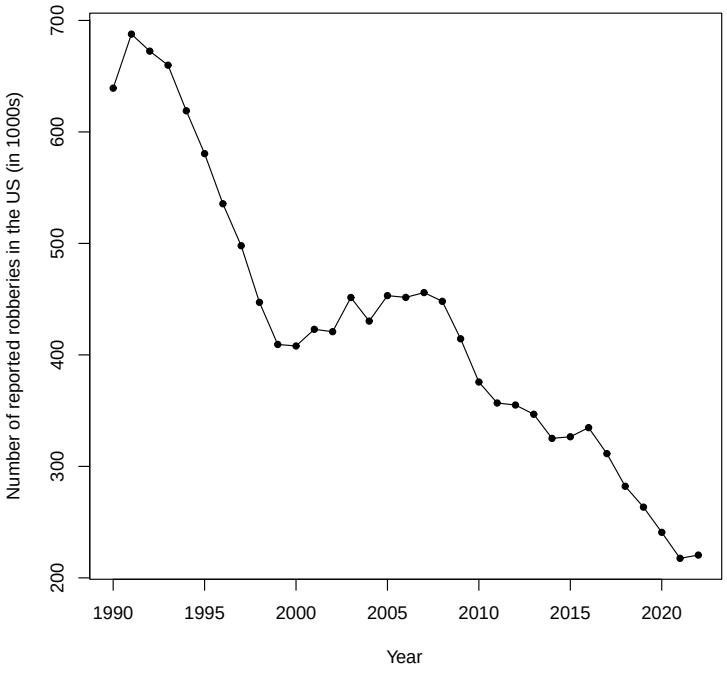
(a) bar graph
(b) scatterplot
(c) line graph
(d) pie chart

7. Federal law requires that tests given to job applicants be directly job related. The Department of Labor believes an employment test called the General Aptitude Test Battery (GATB) should be used. Historical data show scores on the GATB are strongly related to outcomes related to job performance, such as getting a promotion, a raise, positive annual reviews, or an award.

As an instrument to measure job performance, we would say the GATB has

(a) causal variability
(b) accuracy
(c) predictive validity
(d) reliability

8. The data shown in the graph below are the number of reported robberies in the United States (in 1000s) each year from 1990 to 2022 (Source: FBI).



What sentence best describes what we see in the graph?

(a) There is strong evidence of seasonal variation.
(b) There is a clear trend over the 33-year time span.
(c) There are at least two striking deviations from the overall pattern.
(d) All of the above.

9. In the language of clinical trials, what is meant by the term **equipoise**?

(a) Patients must be treated equally throughout the course of the trial.
(b) There is genuine uncertainty about which treatment may benefit a patient the most.
(c) The institutional review board ensures all patients are protected from harm.
(d) Equal (or proportional) randomization should be used in any setting comparing drug or medical interventions.

10. Three hundred children, aged 6-17, will participate in a randomized experiment to compare three painkillers:

- **Group 1:** Ibuprofin
- **Group 2:** Acetaminophen
- **Group 3:** Codeine.

A standard dose of each drug will be administered.

Researchers believe children who participate in athletics may respond differently to the painkillers than those children who do not participate in athletics. If this is true, what experimental design should be used?

(a) randomized complete block design
(b) matched pairs design
(c) completely randomized design

11. In Question 10, no control group was used, that is, there was no group that received a placebo painkiller. If the placebo effect is real, what will likely happen during the experiment?

(a) The benefit of taking the real painkillers will be exaggerated.
(b) More children will be needed to remove the placebo effect as a lurking variable.
(c) None of the group differences will be statistically significant.
(d) Researchers will have to reveal the names of the children.

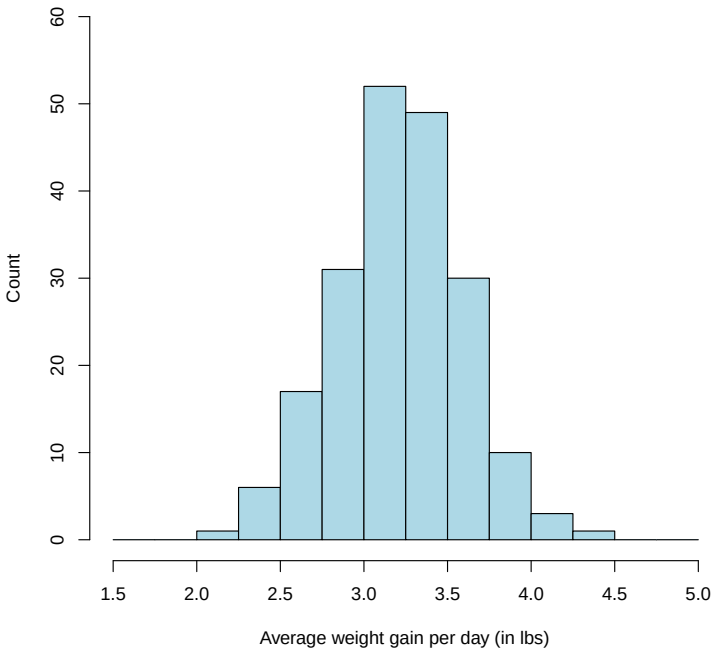
12. Public health officials in the United Kingdom collected data on the counts and rates of maternal deaths for vaginal delivery and Cesarean delivery over a recent 3-year period. Here is a table of the data:

Delivery type	# of deliveries	# of deaths	Death rate per 100,000 deliveries
Vaginal	1,571,000	75	4.8
Cesarean	425,000	73	17.2

A government representative says vaginal deliveries are more dangerous than Cesarean deliveries because a larger number of deaths occurred over this 3-year span (75 versus 73). She is using a measurement which is

(a) not reliable.
(b) misclassified.
(c) invalid.
(d) replicated.

13. A new diet was given to a sample of 200 beef cattle during a test period of 140 days. For each animal, researchers recorded the average weight gain per day (in lbs). A histogram of the 200 observations is shown below:



Which best describes the **shape** of this distribution?

(a) skewed to the left
(b) bimodal
(c) symmetric
(d) skewed to the right

14. Many students sell their own blood plasma to make extra money for college. Suppose you decide to visit a local plasma donor center to donate plasma. Before your plasma can be taken, you are asked to sign a document that describes how your plasma sample will be used, what specifically it will be used for, and any possible risks to you. What is this document asking you for?

(a) institutional review board approval
(b) a confidentiality agreement
(c) a covenant agreement not to take legal action against the donor center
(d) informed consent

15. We talked about the Duke-Potti cancer clinical trial as an example of unethical research behavior, and we watched a 60 Minutes story on it. What exactly did Dr. Potti do that was unethical?

(a) He released confidential information about specific patients to the *New York Times*.
(b) He hacked journal submission web sites to get his papers published without peer review.
(c) He knowingly served on the institutional review board that reviewed his own clinical trial.
(d) Whenever he observed data from a patient whose outcomes did not support his theory, he manipulated the data.

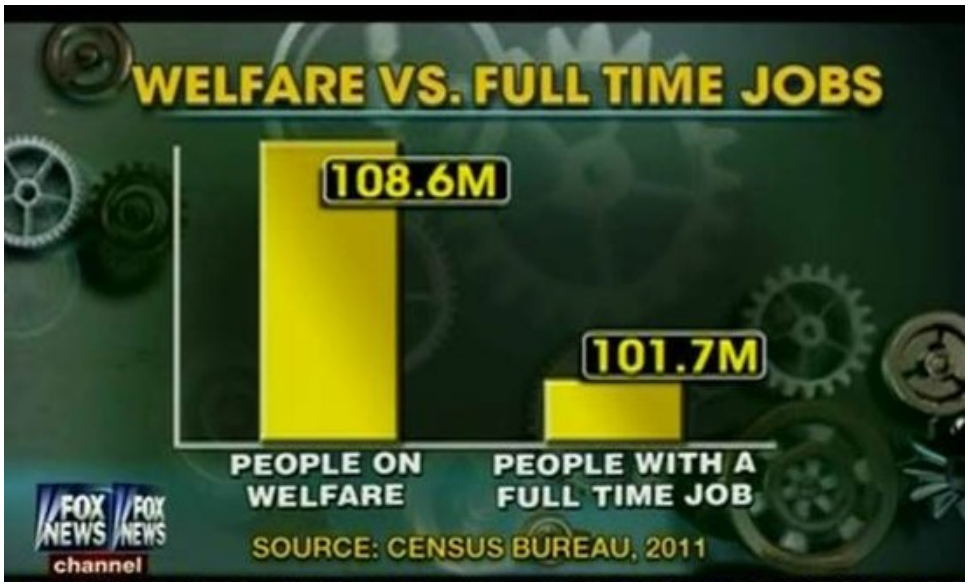
16. At the beginning of the decisive 5th set of the 2017 Australian Open Final between Roger Federer and Rafael Nadal, ESPN displayed the results to the question, *"Who do you think will win?"* Viewers used the internet to vote on this question. At one point, the results posted on the screen were

Federer: 61% Nadal: 44%

What's wrong with these results?

(a) There is no margin of error reported to assess predictive validity.
(b) The numbers do not make sense.
(c) Percentage changes are a more reliable measure.
(d) The number of votes should be reported instead of percentages.

17. In August 2012, Fox News displayed this graph on one of its nightly programs:



Why is this graph misleading?

(a) The viewer cannot tell if the difference between the groups is statistically significant.
(b) Counts of people are quantitative; not categorical. Fox News should have used a stemplot to display these two figures.
(c) The bar height for welfare is about 4 times larger than the height for the full time job group. The difference in counts between the two groups is nowhere close to that.
(d) Fox News did not say if the data in the figure were from an observational study or an experiment.

18. At a recent grant review panel on clinical trials, one panel member (i.e., me) said,

> *"The investigators have yet to obtain approval from their institutional review board."*

What was the main point of my objection?

(a) The authors had not proved how their sample sizes would find statistically significant results.
(b) The authors had not described if their trial would be multi-center or housed all at one common location.
(c) The authors had not provided convincing evidence the subjects would be protected from possible harm.
(d) The authors had not ensured all confounding variables would be controlled.

19. What are the three basic principles of experimental design?

(a) margin of error, confidence, statistical inference
(b) realism, blinding, and equipoise
(c) bias, random error, reliability
(d) randomization, replication, and control

20. Intraocular pressure is the fluid pressure inside the eye. An optometrist has two instruments she uses to measure eye pressure (measured in mm Hg). Five measurements using each instrument are made on a single patient at one time. Here are the results:

Instrument A: 15, 14, 15, 16, 15 Instrument B: 14, 16, 18, 10, 17

Which instrument is more reliable?

(a) Instrument A
(b) Instrument B
(c) The instruments are equally reliable because the average of the 5 measurements is the same for each instrument (15 mm Hg).

21. Kraft's MiO, a liquid water enhancer, comes in a variety of flavors and a few drops added to water gives a zero-calorie flavored water beverage. Researchers want to determine whether those who drink flavored water like the taste of MiO as well as they like pre-flavored water that comes ready to drink. Researchers recruit 50 subjects and set up a comparative experiment with two groups:

- **Group 1:** water + Kraft MiO
- **Group 2:** pre-flavored water.

Each subject will drink the beverage from both groups. The order of which beverage is consumed first is determined at random. Subjects will give a flavor rating for each beverage.

What type of experimental design is this?

(a) completely randomized design
(b) matched pairs design
(c) randomized complete block design

22. A sample of 100 cars with Ford's EcoBoost engine was tested for highway driving. The miles per gallon (mpg) was measured for each car. Below is a stem plot of the observations. The stem is the tens and units digit (e.g., 31). The leaf is the tenths digit (e.g., 0.8).

```
> stem(mpg,scale=2)
    The decimal point is at the |

  30 | 0
  31 | 89
  32 | 5799
  33 | 126899
  34 | 024588
  35 | 0123566789
  36 | 0123344556677788899
  37 | 0011122334456677899
  38 | 01223456789
  39 | 0000345789
  40 | 0123557
  41 | 002
  42 | 1
  43 |
  44 | 9
```

The **mode** of a distribution is the value that appears the most often in the data set. What is the mode for these data?

(a) 36.9
(b) 30.0
(c) 44.9
(d) 39.0

23. I read this statement on a news media web site:

> *"Government has increased medical consultation fees for general practitioners by 100%, state media has reported."*

What does this mean?

(a) Medical consultation fees have increased by 100 dollars per practitioner.
(b) Medical consultation fees have doubled.
(c) This statement cannot be true. It is not possible for a quantity to increase by 100%.
(d) Medical consultation fees have increased by 100 times their previous amount.

24. In class, we talked about a 1994 PhD dissertation whose author claimed,

> *"Every year since 1950, the number of American children gunned down has doubled."*

What was the **main point** of our discussion on this?

(a) Using quantitative summary statistics can distort reality when you are measuring a categorical variable.
(b) We cannot understand the root cause of gun violence unless we use a randomized comparative experiment to control for the effects of lurking variables.
(c) Statistics show that gun violence is steadily getting worse in the United States.
(d) None of the above.

25. Researchers randomly assigned 58 severe-stroke patients to receive either

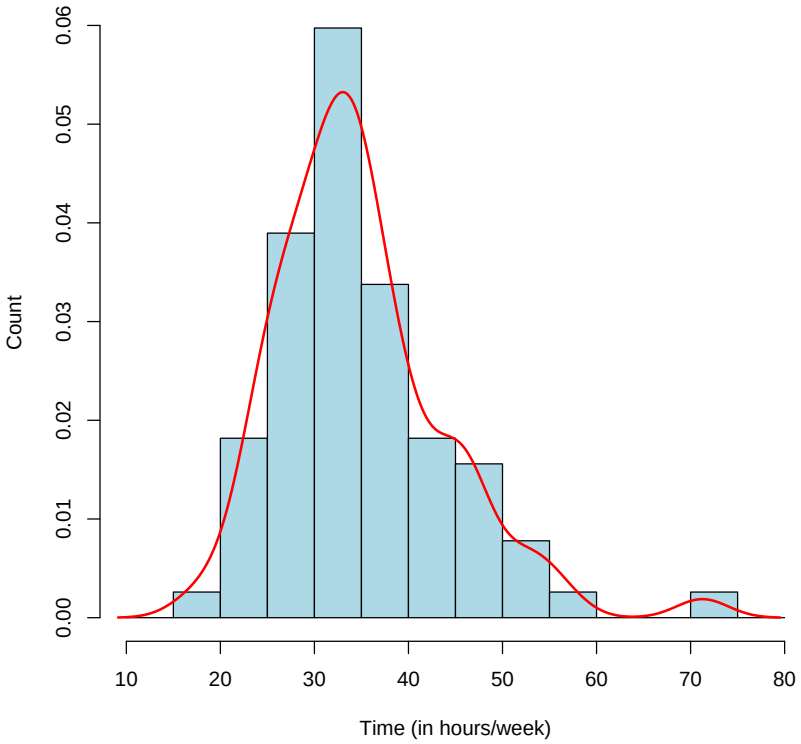
1. **Group 1:** tPA (a standard treatment for stroke)
2. **Group 2:** tPA + blood chilling.

The researchers concluded, "the difference in recovery times between the two treatments was **not** statistically significant." What does this mean?

(a) The experiment was biased because researchers used severe-stroke patients only.
(b) More treatment groups should be used to obtain a statistically significant difference.
(c) The difference between the groups could be explained by random chance.
(d) The sample sizes (29 per treatment) provided differences which are not reliable.

PART 2: SHORT ANSWER. Give a detailed response. Please write clearly.

Twins tend to have lower IQ scores in their early years than non-twins. Psychologists believe this may be explained by the fact that parents do not spend enough time with twins. A study involving a sample of $n = 77$ pairs of twins (aged 5 years and younger) took place in California. Researchers recorded the number of hours parents spent giving attention to their twins during a one-week period. Here is a histogram of these times:



(a) Describe our **four** physical characteristics you see in this histogram.

Answer: Our 4 characteristics are:

- Center: The center of the distribution is around 35 hours/week.
- Variability: The times range from about 15-60 hours/week (excluding the outlier).
- Shape: The distribution is slightly skewed to the right (high) side.
- Deviations: There is one outlier around 70-75 hours/week.

(b) Does it make sense to think about a **population density curve** in this example?

Answer: Yes, it does. The histogram shows the distribution of the times for the sample of 77 twins. A population density curve would give the distribution of the times for the entire population, that is, the population of all twin pairs (aged 5 years and younger), or perhaps all twin pairs in California (whatever you envision the population to be).

PART 3: EXTRA CREDIT. Give a detailed response. Please write clearly.

We have used the term "innumeracy" in class as described in Joel Best's book *Damned Lies and Statistics*.

(a) Explain what innumeracy means.
(b) Give a real-life example where innumeracy might arise. Your example can be one we discussed in class or you can make up your own.
(c) What is Dr. Best's book about?

Answers:
(a) "Innumeracy" is a made up word by Dr. Best, but it is meant to serve the same purpose as "illiteracy" does when it comes to reading. Innumeracy is the state of not being comfortable with basic mathematical concepts, including not understanding numbers. It also has to do with people not checking numbers/math/calculations for themselves and just taking another person's word for it. This is often done in the media.

(b) We talked about a lot of examples in class:

- Louisiana biology teacher example
- NYT headline claiming that over 100 percent of Americans would fail a test
- Doubling of children "gunned down" every year example (maybe the worst)
- Examples claiming a percentage decrease can be larger than 100 percent
- 31,000,000 Americans hungry example (the definition of "hunger" was debatable)
- The Organic Gardening example claiming the US Interstate Highway system spanned 3.9 million miles (it's much less than that).

(c) Dr. Best's book is about how the media abuses numbers and statistics to bring attention to social issues and other agendas.