Namecheq Dat Stat

1. We wonder if the new Kanye West single is as significantly different as advertised. On average, Kanye West songs have 10.1 beats per minute, with a variance of 7.9. This song has a mean 11.1 beats per minute, with a standard deviation of .80.

   One sample z-test

2. There is a violent discussion over the distinction between stealing versus sampling old hits. Data is then collected as to whether people believe they steal (n = 20) versus sample classics (n = 48) or whether they believe there isn't a distinction (n = 100).

   Chi-square test

3. A researcher examines differences in the frequency of using colloquial words in people who have versus have not completed a college degree, accounting for region and socioeconomic status.

   If normal, Paired t-test (2 sample)

4. MF Doom and Aerosmith decide to combine their talents to produce a new musical genre ("Aerodoom"). We believe approximately half of the fans of this new genre listened to either MF Doom or Aerosmith before they became Aerodoom fans. We examine Aerodoom fans and determine how many people listened to MF Doom (n = 199), Aerosmith (n = 200) or neither (n = 25) before becoming Aerodoom fans.

   Chi-square goodness of fit

5. We investigate the number of times people listen to classical music, depending on whether they drive (n = 40, mean = 38.14, SD = 4.87), take the bus (n = 40, mean = 24.41, SD = 16.83), ride a bike (n = 40, mean = 28.4, SD = 3.2) or walk (n = 20, mean = 14.5, SD = 2.2).

   HOV: $\frac{16.83^2}{2.2^2}$ = 58.52, Kruskal-Wallis

6. We examine the number of songs people belt out in their cars in America (mean = 5.45, SD = 1.24). In Europe and other continents, people generally sing 1.5 songs on average.

   One-sample t-test

7. We wonder how many times Malcolm McLaren stole from the Sex Pistols. We decide to ask each band member. We believe, however, that their decision may be confounded by their current financial situation. We therefore ask each member before versus after receiving a large sum of money. Sadly, S. Vicious dies before the end of our study.

   If normal/HOV, independent t-test (2 sample)

8. We examine sex differences in the number of times people order martinis. Females order 1.4 martini drinks, with a SD of 4.5. Males order 2.4 martini drinks, with a SD of 14.8.

HOV: $\frac{14.8^2}{4.5^2}$ = 10.82, Mann-Whitney U

9. We are fascinated in the effects of music preference on the number of times people begin to dance randomly. We sample people who listen to rock (n = 42, mean = 14.56, SD = 4.23), rap (n = 42, mean = 14.44, SD = 2.34) or country music (n = 42, mean =14.52, SD = 2.44).

HOV: $\frac{4.23^2}{2.34^2}$ = 3.27, One-way ANOVA

10. A researcher examines how often people dream as well as how often they laugh weekly. She finds people dream 7.5 times a week, on average, with an SD of 1.4.

Correlation

11. We believe that travel may affect interest in other cultures. We ask people how often they leave the country as well as how often they pursue opportunities to learn more about other cultures.

Regression (Travel "affects" interest)

12. A statistician's wife dies tragically in a car accident. Others believe he is crazy, but he believes she knew what was going to happen. On average, she would smile for 1.4 seconds, with a standard deviation of .40. The last time she saw him before the car accident, however, she smiled for 4.5 seconds.

One sample z-test

13. We investigate how often people rank themselves as smartest in their intellectual circles relative to whether they've successfully read *Ulysses*. We control for education level, because college students may feel more pressured to read *Ulysses.*

Wilcoxon T

14. We examine the relationship between political party and voter registration in 500 college students. We find that 60% of Republicans (n = 225) are registered, whereas 50% of Democrats (n = 275) are.

Chi-square Contingency

15. We investigate how much undergraduates hate statistics after taking required coursework in Psychology (n = 100), Statistics (n = 100) or Math (n = 100), from a scale of 0 – 5.

Kruskal-Wallis

16. Public opinion believes 70% of celebrities advertise but do not actually commit to rigorous dieting. We collect data and find 150 advertise, but do not diet; 240 do not advertise, but do diet; and 140 do not advertise or diet.

Chi-square goodness of fit

17. We investigate the effects of ketamine on baby crying. We treat some children, who cry 0.4 times, on average, with a SD of 1.4. The other group of children, left untreated, cry a mean 5.4 times, with a SD of 2.3.

HOV = $\frac{2.3^2}{1.4^2}$ = 2.70, Independent t-test (2 sample)

18. We believe mood (happy, neutral, upset) may affect whether people enjoy *Happy Feet*. We find 100 people enjoyed the movie when happy, 50 when neutral, 20 when upset. In contrast, 10 people did not appreciate *Happy Feet* when happy, 40 when neutral, 70 when upset.

Chi-square contingency

19. We examine effects of taking a test on anxiety levels in 29 psychology students, sampling students before versus after an exam. Our skew, however, is greater than 12, as is our kurtosis.

Wilcoxon T

20. How does one determine if just as many people mourned Anna Nicole as Aaliyah and Lisa “Left Eye” Lopes?

Chi-square Test

21. We believe that marriage may affect how often people compromise under social pressure. We distribute a survey to single and married people, determining how often subjects feel they compromise (1 = few times, 5 = many times).

Mann-Whitney U

22. A researcher investigates how often people compute non-parametric statistics, if their data has violated assumptions of parametric tests. He investigates people whose data are not normally distributed (n = 400, mean = 40, SD = 4), do not have homogenous variances (n = 400, mean = 70, SD = 3) or are ordinal (n = 500, mean = 39, SD = 3).

HOV = $\frac{4^2}{3^2}$ = 1.78, One-way ANOVA

23. We examine the number of gladiator movies made by producers who have versus have not spent time in Italy. We control for age and socioeconomic status, which may affect the ability to travel.

If normal, paired t-test (2 sample)

24. A researcher investigates the relationship between the yearly number of electronic artists in Britain (Mean = 204.98, SD = 28.47) and the yearly number of new electronic genres (Mean = 3.44, SD = 1.08).

Correlation

It's More Fun to Compute

1. A researcher investigates the relationship between diet coke and ulcers in a sample of 100 college girls. She finds a positive correlation (r = .53).

    a. How "good" is this correlation, according to rules of thumb in this class?
    .53 → Modest/OK

    b. How much variability in diet coke is actually explained? Ulcers?
    $.53^2$ = 0.28. 28%. It will be the same for both variables -> correlation does not imply causation.

    c. Is this correlation significant?
    r_{crit} (df = 100-2) = .195 -> Yes.

    d. State your conclusions.
    There is a modest positive association between drinking diet coke and getting ulcers. The more one drinks diet coke, the more likely one will get an ulcer.

2. A scientist believes the number of times a bonobo chimp listens to James Brown can affect his abilities to attract mates (N = 30). She finds the following relationship, which explains 84.46% of the variance.

    $Y = 2.4123x + 1.09$

    $S_{x|y} = .8738$

    a. State the null and alternative hypotheses. Be explicit – what do these hypotheses mean?

    H0: The number of times a bonobo chimp listens to James Brown can not be used to predict his abilities to attract a mate.How "good" is the correlation.
    H1: The number of times a bonobo chimp listens to James Brown can be used to predict his abilities to attract a mate. (Implies causation)

    b. How "good" is the correlation coefficient between these variables, according to rules of thumb in this class?
    Sqrt(.8446) → 0.92. Excellent/Strong correlation coefficient.

    c. If a male listened to James Brown 5 times, how many mates would he be predicted to attract?

    $Y = 2.4123(5) + 1.09$ → 13.15

d. What is the probability of attracting 15 mates, if a male listens to James Brown 8 times? Could this probability be due to chance alone?
Y = 2.4123(8) + 1.09 → 20.3884 (predicted number of mates)
$\frac{20.39 - 15}{.8738}$ = 6.17 → z-score.
Probability of 6.17 → Not on table → p < 0.0001

3. We become fascinated in the relationship between listening to hip hop and hating statistics in a sample of students (r = .78, N = 46).

   a. Is this correlation significant?
   r_{crit} (df = 78-2) = .217-> Yes.

   b. What are the possible conclusions/inferences we can make from this relationship?
   Listening to hiphop causes hatred of statistics.
   Hatred of statistics causes listening to hiphop.
   Some other variable (i.e. Shoe size) causes both hatred of statistics and listening to hiphop.
   Type I Error.

4. A therapist examines the relationship between consuming hallucinogens and writing esoteric poetry in 450 high school students. He records the number of hallucinogens and poems written for each student. He finds a correlation co-efficient of .52, y-intercept of 2.485 and a slope of .8746.
   a. Is this a significant correlation?
   r_{crit} (df = 450-2) = .098 -> Yes.

   b. How much variance is explained?
   $.52^2$ = .2704. 27%.

   c. One student who has taken 7 hallucinogens writes 20 poems. What is the probability of this occurring due to chance alone?
   Y = .8746(7) + 2.485 → 8.6072 (y-expected).
   I forgot to include the SEE (obviously won't be a problem on the exam, you'll be given all the info) → Let's say it's 2.5812.
   $\frac{20 - 8.6072}{2.5812}$ = 4.413761 → Off the chart, so p < 0.0001.

5. There's a UW INSTRUCTOR MASSACRE!

	Statistics	Philosophy	Business
Number Killed:	69	20	41

a. What statistic should you compute? Justify your answer.
No other information given except observed distribution → Regular chi-square.

b. Compute, state null/alternative hypotheses and conclusions.
H0: The number of UW instructors murdered is evenly distributed among departments.
H1: The number of UW instructors murdered is not evenly distributed among departments.

Observed	Expected	O-E	O-E2	O-E2/E
69	43.3333333	25.6667	658.7778	15.2026
20	43.3333333	-23.3333	544.4444	12.5641
41	43.3333333	-2.3333	5.4444	0.1256
130	43.3333333			**27.8923**

Chi-square crit (3-1) = 5.99. Reject null hypothesis. The number of UW instructors is not evenly distributed among departments.

6. We are among a series of massacres on college campuses. Generally, 50% of instructors massacred are businessmen, 25% philosophers, 25% statisticians.

a. What statistic should you compute? Is this more or less powerful than (or the same as) the statistic in #5? Justify your reason.
Chi-square goodness of fit.

b. Compute, state null/alternative hypotheses and conclusions. Be specific to the story – what do your conclusions mean in relation to UW (vs other college massacres)?
H0: The UW massacre will be distributed like massacres on other college campuses.
H1: The UW massacre will not be distributed like massacres on other college campuses.

Observed	Proportion	Expected	O-E	O-E2	O-E2/E
69	(130*.50)	65	4	16	0.2461
20	(130*.25)	32.5	-12.5	156.25	4.8077
41	(130*.25)	32.5	8.5	72.25	2.2231
130					**7.2769**

Chi-square crit (3-1) = 5.99. Reject null hypothesis. The number of UW instructors is not distributed like massacres on other college campuses.

7. We examine the relationship between taking statistics and murdering. Compute the appropriate stat, state null/alternative hypotheses and conclusions.

	Statistics	No Statistics	
Murderin'	60	40	**100**
No Murderin'	20	15	**35**
	80	**55**	**135**

Expected

	Statistics	No Statistics
Murderin'	(80*100)/135	(55*100)/135
No Murderin'	(80*35)/135	(55*35)/135

Observed	Expected	O-E	O-E2	O-E2/E
60	59.2593	0.740741	0.5487	0.0093
20	20.7407	-0.740741	0.5487	0.0265
40	40.7407	-0.740741	0.5487	0.0135
15	14.2593	0.740741	0.5487	0.0384
				0.087662

Chi-square crit (2-1*2-1) = 3.84. Accept null hypothesis.

8. Calculate the following, using the appropriate chi-square. State conclusions.

Chi-square contigency

	Death	No Death	
Bleeding	90	50	140
Some Bleeding	42	48	90
No Bleeding	10	42	52
	142	140	
			282

Expected

	Death	No Death
Bleeding	142*140/282	140*140/282
Some Bleeding	142*90/282	140*90/282
No Bleeding	142*52/282	140*52/282

Observed	Expected	O-E	O-E2	O-E2/E
90	70.4965	19.5036	380.3883	5.3959
42	45.3191	-3.3191	11.0168	0.2431
10	26.1844	-16.1844	261.9347	10.0035
50	69.5036	-19.5035	380.3883	5.4729
48	44.68085	3.3191	11.0168	0.2466
42	25.8156	16.1844	261.9347	10.1464
				31.50828

Chi-square crit (3-1*2-1) = 5.99. Reject null hypothesis.

Chi-square test

	Albums Sold
Bald Britney	145
Mobb Deep	300
Nirvana	150
Happy Hardcore	50
Mozart	130

Observed	Expected	O-E	O-E2	O-E2/E
145	155	-10	100	0.6452
300	155	145	21025	135.6452
150	155	-5	25	0.1613
50	155	-105	11025	71.1290
130	155	-25	625	4.0322
775				**211.6129**

Chi-square crit (5-1) = 9.49. Reject null hypothesis.

Chi-square contigency

	Halo	Zelda	Street Fighter	
American Idol	5	15	20	**40**
Grey's Anatomy	1	1	3	**5**
The Wire	80	60	42	**182**
	86	**76**	**65**	**227**

Expected

	Halo	Zelda	Street Fighter
American Idol	(86*40)/227	(76*40)/227	(65*40)/227
Grey's Anatomy	(86*5)/227	(76*5)/227	(65*5)/227
The Wire	(86*182)/227	(76*182)/227	(65*182)/227

Observed	Expected	O-E	O-E2	O-E2/E
5	15.1542	-10.1542	103.1075	6.803894
15	13.3921	1.60793	2.585437	0.193057
20	11.4537	8.546256	73.03848	6.376821
1	1.89427	-0.89427	0.799724	0.42218
1	1.67401	-0.67401	0.454288	0.271377
3	1.431718	1.568282	2.459508	1.717872
80	68.95154	11.04846	122.0684	1.770351
60	60.93392	-0.93392	0.872208	0.014314
42	52.11454	-10.1145	102.3039	1.963058
				19.53293

Chi-square crit (3-1*3-1) = 9.49. Reject null hypothesis

9. We believe the relationship between taking statistics and murdering is mediated by a student's final GPA. We examine GPAs from murdering versus non-murdering statistics students. Compute the appropriate non-parametric test.

GPA

Murdering	2.1	3.8	1.9	2.4	1.4	3.0	2.0
Non-Murdering	4.0	1.8	2.7	2.9	1.9	3.1	4.0

Mann-Whitney U.

Groups	N	Mean Rank	Sum of Ranks
Murdering	7	6.36	44.50
Non-Murdering	7	8.64	60.50
Total	14		

U = 16.5000
Ucrit = (2-tailed, alpha = .05) → 8. Accept null hypothesis.

10. How could we increase our power for the previous study above? Re-compute the above problem using each suggestion.
    1. Meet parametric assumptions (You WON'T have to compute this for the exam)-> Independent t-test. -1.205, df = 12. (tcrit = 2.179 → Accept null hypothesis).
    2. Use a paired design ->

	N	Mean Rank	Sum of Ranks
Negative Ranks	1	6.50	**6.50**
Positive Ranks	6	3.58	21.50
Ties	0		
Total	7		

Tcrit = 2. Accept null hypothesis.

3. Increase alpha → 11. Accept null hypothesis.
    a. One-tailed (Make it alpha .05 one tailed)
    b. Two-tailed (Make it alpha .10 two-tailed)
4. Increase sample size (don't worry about this)

11. Compute the following data – first as a Wilcoxon, then as a Mann-Whitney. (Assuming 2-tailed, alpha = .05)

a.

4.00	10.00
3.00	8.00
2.00	8.00
7.00	9.00
3.00	9.00
4.00	15.00
2.00	7.00
5.00	8.00
4.00	12.00
6.00	7.00

Wilcoxon

	N	Mean Rank	Sum of Ranks
Negative Ranks	0	.00	.00
Positive Ranks	10	5.50	55.00
Ties	0		
Total	10		

Wilcoxon Tcrit = 8 -> Reject null hypothesis.

Mann-Whitney U

Groups	N	Mean Rank	Sum of Ranks
1.00	10	5.60	56.00
2.00	10	15.40	154.00
Total	20		

U = 1.0.
Ucrit = 23 -> Reject null hypothesis.

b.

3.82	6.52
10.10	4.73
4.15	7.28
5.11	7.74
13.50	11.85
14.84	9.56
9.75	11.05
13.53	6.34
18.74	4.68

Wilcoxon

	N	Mean Rank	Sum of Ranks
Negative Ranks	5	6.40	32.00
Positive Ranks	4	3.25	**13.00**
Ties	0		
Total	9		

Wilcoxon Tcrit 9 = 6 -> Accept null hypothesis.

Mann-Whitney U

Groups	N	Mean Rank	Sum of Ranks
1.00	9	10.78	97.00
2.00	9	8.22	74.00
Total	18		

U = 29.00
Ucrit = 17 -> Accept null hypothesis.

c.

12.74	2.00
10.67	7.00
10.96	3.00
10.97	8.00
11.77	.00

Wilcoxon

	N	Mean Rank	Sum of Ranks
Negative Ranks	5	3.00	15.00
Positive Ranks	0	.00	**.00**
Ties	0		
Total	5		

Wilcoxon Tcrit 5 = 0 -> Reject null hypothesis.

Mann-Whitney U

Groups	N	Mean Rank	Sum of Ranks
1.00	5	8.00	40.00
2.00	5	3.00	15.00
Total	10		

U = 0.
Ucrit = 2.

d.

1	1
4	3
3	2
4	5
7	7
6	5
6	1
0	0
7	4

Wilcoxon

	N	Mean Rank	Sum of Ranks
Negative Ranks	5	3.70	18.50
Positive Ranks	1	2.50	**2.50**
Ties	3		
Total	9		

Wilcoxon Tcrit 9 = 6 -> Reject null hypothesis.

Mann-Whitney U

Groups	N	Mean Rank	Sum of Ranks
1.00	9	10.67	96.00
2.00	9	8.33	75.00
Total	18		

U = 30.
Ucrit = 17 -> Accept null hypothesis.

12. Calculate K-W for the following data. Also – state why you might prefer K-W to ANOVA for these columns/groups.

a.

.95	3.37	7.65
-1.92	4.42	11.09
.84	5.79	3.72
1.52	4.52	7.42
-.24	3.46	12.14

HOV violation

Groups	N	Mean Rank
1.00	5	3.00
2.00	5	8.60
3.00	5	12.40
Total	15	

K-W: 11.180.
Chi-square/K-crit (3-1) = 5.99. Reject null hypothesis.

b.
9.00	6.00	100.00
8.00	5.78	25.00
9.00	6.02	25.00
8.00	6.00	14.00

HOV violation

Groups	N	Mean Rank
1.00	4	6.50
2.00	4	2.50
3.00	4	10.50
Total	12	

K-W: 9.986.
Chi-square/K-crit (3-1) = 5.99. Reject null hypothesis.

c.
9.90	12.47	15.29	19.52
12.35	12.24	15.41	15.79
13.66	13.74	16.40	45.40
-2.08	8.88	13.84	12.22
9.19	11.68	15.56	5.62

HOV/normality?

Groups	N	Mean Rank
1.00	5	6.00
2.00	5	7.80
3.00	5	15.20
4.00	5	13.00
Total	20	

K-W: 7.983
Chi-square/K-crit (4-1) = 7.82. Reject null hypothesis.