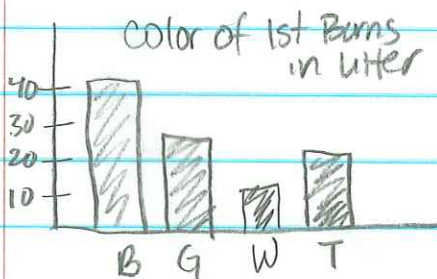


AP Stat

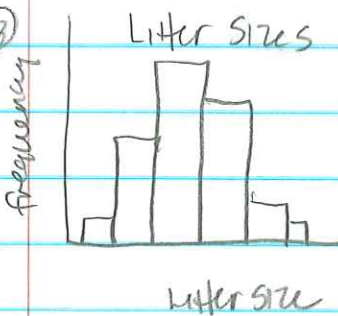
Midterm Review

① categorical = color quantitative = amount in liter

②



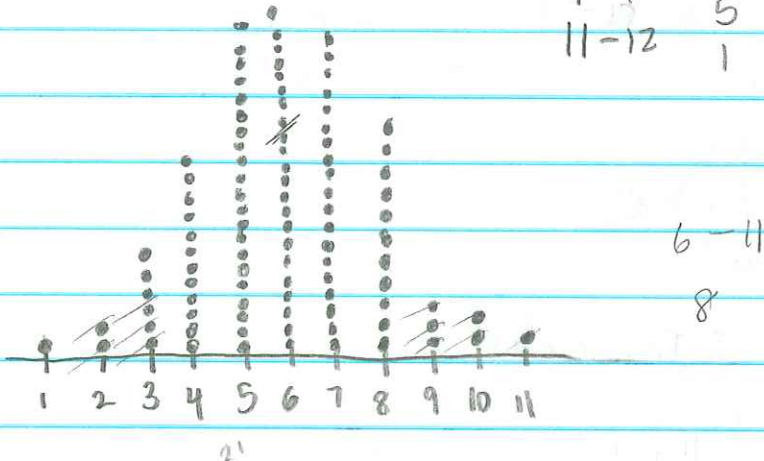
③



$$\text{range} = 11 - 1 = 10 \quad 10 \div 6 = 1.$$

1-2	3
3-4	18
5-6	43
7-8	30
9-10	5
11-12	1

④



$$6 - 11$$
$$8$$

⑤ Does not appear to be any outliers - possibly 1 & 11

⑥ Histogram appears symmetric

⑦ mean = 5.88 the mean is not resistant to outliers

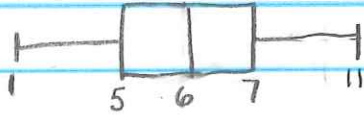
⑧ median = 6 the median is resistant to outliers

⑨ range = $11 - 1 = 10$

⑩ min = 1 Q1 = 5 Q3 = 7
max = 11 Q2 = 6

11) $IOR = 7 - 5 = 2$

12)

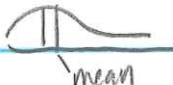


13) standard deviation is not resistant
st dev = 1.82

14) Area = 1

15) median \rightarrow equal areas

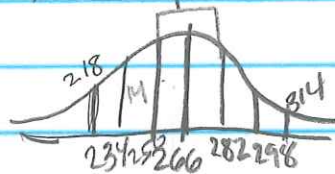
16) mean \rightarrow balancing point

17)  mean further right than median

18) $\bar{x} \rightarrow$ sample mean $\mu \Rightarrow$ population mean

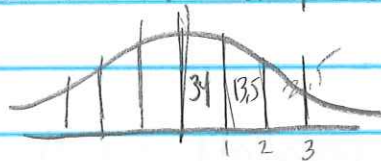
19) $S_x \rightarrow$ sample st. dev $\sigma \Rightarrow$ population st. dev

21)



22) 68-95-99.7 Empirical rule for normal curve

23)



$x > 282$ days

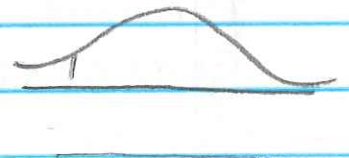
24) $218 < x < 314$ days

25) $x < 234$ days

26) .15 percentile $z = \frac{218 - 266}{16} = -3$

27) 97.5 percentile

28) 266 is 50th percentile



$$(29) \quad z = \frac{x - \mu}{SD} \quad z = \frac{257 - 266}{16} = \boxed{-.5625 = z}$$

$$(30) \quad x < 257 \\ z < -.5625 \Rightarrow p = .2877 \Rightarrow \boxed{28.77\%} \text{ (Chart)}$$

$$(31) \quad x > 280 \\ \text{normalcdf}(280, \infty, 266, 16) = .191 \Rightarrow \boxed{19.7\%} \text{ (Calc)}$$

$$(32) \quad 260 < x < 270 \\ \text{normalcdf}(260, 270, 266, 16) = .2449 \Rightarrow \boxed{24.49\%}$$

$$(34) \quad \text{inv.norm}(.9, 0, 1) \quad z = 1.28 = \frac{x - \mu}{SD} \quad 1.28 = \frac{x - 266}{16} \\ 20.5 = x - 266 \\ \boxed{x > 286.5 \text{ days}}$$

$$(35) \quad \text{inv.norm}(.25, 0, 1) \quad z = -.674 = \frac{x - 266}{16}$$

$$\boxed{x < 255 \text{ days}} \quad -10.79 = x - 266 \\ 255 = x$$

$$* (36) \quad 261.9 - 270.1$$

(37) yes

	4th	7th
38	0	x
	852	1 2588
	9865200	2 0334578
	97655521	3 00x33356
	20	4

Key: 0|1=1
0|4=40

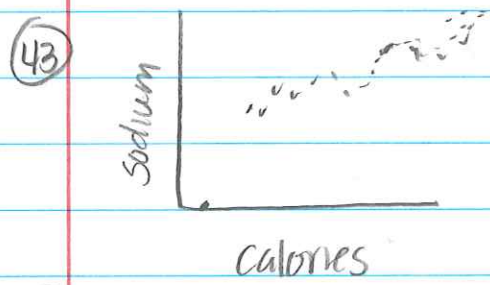
39) 7th median = 23
4th median = 30 ← reading scores better. resid $y - \hat{y}$

40) 7th mode 33
4th mode 20, 35
1QR $Q1 = 19 - 32 - 19 = 13$
 $Q3 = 32$

1.5 IQR = 19.5 $Q1 - 19.5 = .5$

41) 1 is not an outlier Based on 1.5 IQR rule

42) modified box plot is a boxplot that takes outliers out and notes them w/ points not included in Box plot. It more accurately represents the dispersion of data



44) response var is Sodium

45) explanatory var is Calories

46) upward positive direction

47) Linear $\hat{y} = -85.41 + 3.1x$

48) strong linear w/ $r = .919$ $\text{sd} = -85.41 + 3.1 \text{ cal.}$

49) (108, 149) has a large residual → outlier

50) $\hat{y}_2 = 24.4 + 2.45x$ $r = .95$

w/o outlier, equation changes significantly
so yes, influential

51) $r_1 = .91$

52) $r_2 = .95$

53) strong linear

54) Changing units will not change correlation r .

55) highest possible = 1

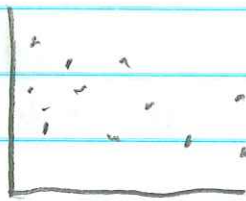
56) lowest is zero

57) correlation applies to quantitative variables linear

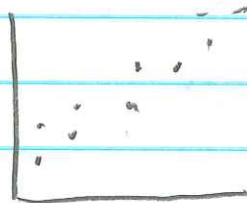
58) Correlation is not resistant to outliers

59) A high correlation does not necess. mean causation
it shows a high association

60)



60)



62) $\hat{y} = -85.41 + 3.1x$

$Sod = -85.51 + 3.1(\text{calories})$

$\frac{\Delta y}{\Delta x}$

$\frac{\Delta Sod}{\Delta cal}$

63) Slope = 3.1: as calories go up sodium goes up at about 3 units per calorie

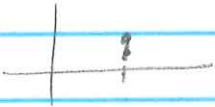
64) $Sod = -85.51 + 3.1(155)$
 $Sod = 394.9$ or 395

65) $Sod = -85.51 + 3.1(315)$
 $Sod = 983.99$ or 984

- (66) #64 interpolation - within data
 #65 extrapolation - outside data makes unreliable

(67) $res = y - \hat{y}$ $\hat{y} = -85.51 + 3.1(180)$
 $\hat{y} = 472.5$

$res = 500 - 472.5$
 $res = 27.5$



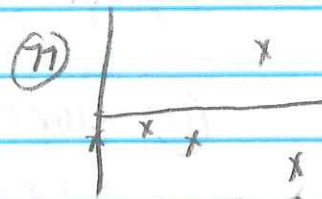
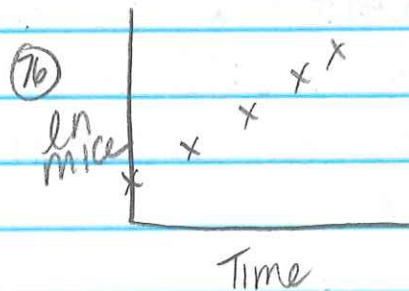
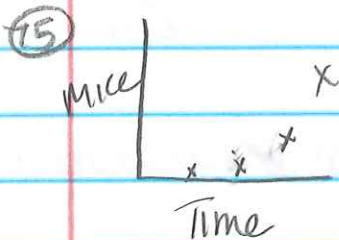
- (68) $\bar{x} = 159.4$ $SD = 23.1$ Cal.
 (69) $\bar{y} = 416.56$ $SD =$ Sod.
 (70) \rightarrow

(71) $r^2 = .844$

(72) 84.4% of variation is explained by the LSRL
 as cal + sodium are related

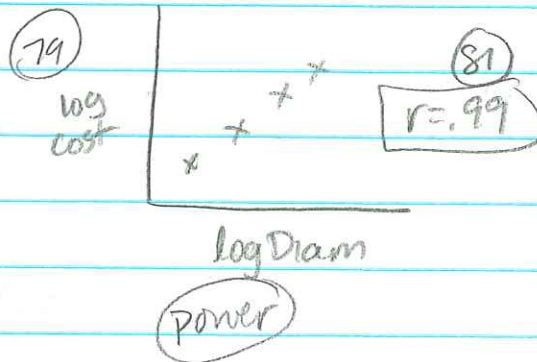
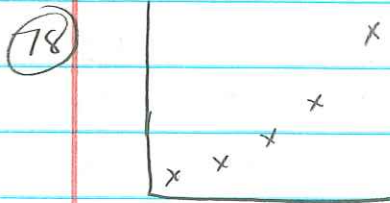
(73) if residuals are evenly distributed above + below
 the $y=0$ line (no discernable pattern)

(74) Examine the r/r^2 and residual plot

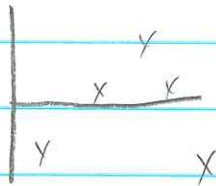


$r = .999$

L1
L5



80 Residual



power

$$\log \text{cost} = -1.02 + 2.017 \log \text{diam}$$

82 All response to #66

83 \uparrow

84 — unaccounted for variable that may effect x/y
 Lurking variable is a type of extraneous variable, its effect on the response variable can not be distinguished from one or more of the explanatory variables

85 H "smooths out outliers effect on correlation"

86 Causation — one variable influences another
 Alcohol causes lab rats temp to go up

87 Common response — something that impacts 2 things
 Ex) Intelligence effects both SAT Math + Verbal scores

* 88 confounding

89 — skip

(90)	NS	S-ng	Sm	DNCHS	CHS	1-3	4+
	357	156	312	183	225	200	186
	<u>825</u>	<u>825</u>	<u>825</u>	<u>825</u>	<u>825</u>	<u>825</u>	<u>825</u>

(91) .38

(94) $\frac{113}{312} = .36$

(92) $\frac{91}{357} = .27$

(95) It appears that of all smokers the highest % didn't complete HS

(93) $\frac{63}{180} = .39$

(96) Simps Paradox - $\uparrow \downarrow$ Ex) Homer Simpson loves donuts + beer yet never gains weight

(97) observational study - does not impose a treatment to compare.

Experiment - treatment is imposed + compared. has controls

(98) V. Resp. sample - a sample composed of volunteers

(99) sample is subset of population that represents the population being examined/studied

(100) Convenience sampling may not represent the pop since units are selected out of convenience (ease)

(101) SRS = Simple Random Sample
Every sample size n has equal chance being selected

- 102 Ex. # each town alphabetically. use random digit table or technol. to select 4 #'s randomly from set.
Ex - names in hat, randomly select 4.
- 103 Stratified - get representatives from each "group" to rep the pop - Layers
- 104 Cluster - smaller "groups" selected that rep. population
- 105 undercoverage - not enough represented.
- 106 non response - not getting responses can lead to bias
B/c those responding may be biased
- 107 Lying
- 108
- 109
- 110 Subjects are people - they are experimental units
- 111 three LEVELS
- 112 Placebo effect is having the response the actual drug entails / or treatment w/o actually getting the treatment but thinking you did → Brain actually may produce chemical so Body responds like it got it
- 113 control group allows for comparison
controls lurking variables
- 117 Block Design - control possible confounding variables at the beginning of an experiment

(121) Independence
The probability of an event does not effect the prob
of another

(122) Sample space is H or T

(123) 0, 1, 2, 3

(124) HHH, HHT, HTH, THH, TTH, THT, HTT, TTT

(124) 12

(127) $10 \cdot 10 \cdot 10 \cdot 10 = 10^4 = 10000$

(128) $10 \cdot 9 \cdot 8 \cdot 7 = 5040$

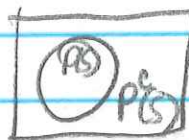
(129)

(130) 0 and 1

(131) 1

(132) 1

(133) $1 - P = \text{comp} \rightarrow \text{failure}$



(134) disjoint - not can be same thing - not connected

Red or Black Card

Ace or King



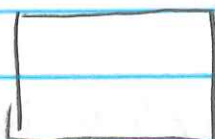
(135) .1

(136) $P(R) = .2$ $P(G) = .1$

$P(R \text{ or } G) = .2 + .1 = .3$

(137) yellow + orange - can't be both $p = 0$

(138)



Bre Beat Erica .09

Erica Beat Bre .08

$(.09)(.92) = .083$

Bre .92

(139) Assumption that Bre Beats Erica swimming .92
(not valid)

(140) 7 or 11

5,2 3,4 1,6 5,6
2,5 4,3 6,1 6,5

$\frac{4}{36}$

(141) 11 33 55 $\frac{6}{36}$
22 44 66

(142) $\frac{8}{36} \cdot \frac{6}{36} = \frac{48}{1296} = .037$

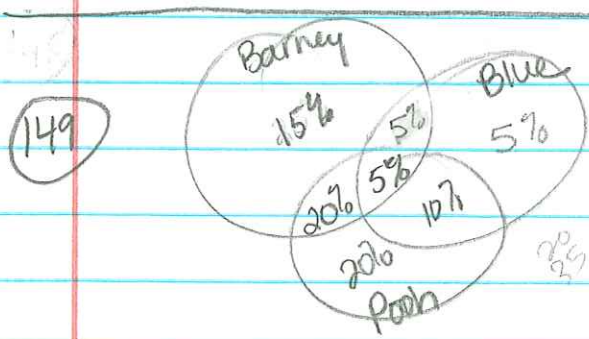
* (143) Rolls are independent - valid

(144) zero

(145) Both 00 sum prob.

(146) 00 Both at same time common outcomes

(147) Check $P(A \text{ and } B) \neq P(A)P(B)$ | OR | $P(B|A) = P(B)$



(148) $P(A+B) = P(A)P(B)$

$A = \text{SMOKING} \quad B = \text{HS}$

$P(A) = \frac{312}{825} = .378$

$P(B) = \frac{214}{825} = .26$

$\frac{113}{825} = .14 \neq .098$

②

$$\pi(10)^2 = 314.16 \quad S_9 = 24^2 = 576$$
$$\frac{314.16}{576} = 54.5\%$$

(151) A = Heart Event A & B not disjoint
B = King King can be heart
C = spade

(152) Yes knowing king or heart does not influence prob of the other

(153) Yes, A Heart can't be a spade

(154) No. If I tell you it's a heart then you know it's not a spade

(155) \cup union (or)

(156) \cap intersection (and)