

Hypothesis Testing

Statistical procedures for addressing research questions involves formulating a concise statement of the hypothesis to be tested.

The hypothesis to be tested is referred to as the *null hypothesis* (abbreviated H_0) because it is a statement of *no difference*.

Hypothesis testing starts with the assumption that the null hypothesis is true... that there is/are no difference(s).

Along with the null hypothesis we must also state an alternate hypothesis (abbreviated H_a).

The alternate hypothesis is a statement that a difference exists.

If a null hypothesis is rejected, then we *tentatively* accept the alternate hypothesis and conclude that there is a difference.

Why is the null hypothesis the one that is tested?

Think about it this way: we only have to find one instance in which the null hypothesis is not true (false) in order to be able to reject it.

Conversely, we would have to continue to test the alternate hypothesis in order to be able to accept it.

In other words we would have to test all possibilities since the alternate hypothesis can only be ***proven*** correct if *all* possible tests are performed.

The moral of the story:

It is easier to prove a null hypothesis incorrect than to prove an alternate hypothesis correct.

Moral number B:

With statistics we never **PROVE** anything. The best we can do is *tentatively accept* a result.

Example Hypotheses:

H_0 : There is no difference in weight between the two groups.

H_a : There is a difference in weight between the two groups.

IMPORTANT:

In ***ALL*** cases, if your calculated probability or probability range is ***less*** than 0.05, then you ***REJECT*** the null hypothesis.

If your calculated probability or probability range is ***greater*** than 0.05, then you ***ACCEPT*** the null hypothesis.

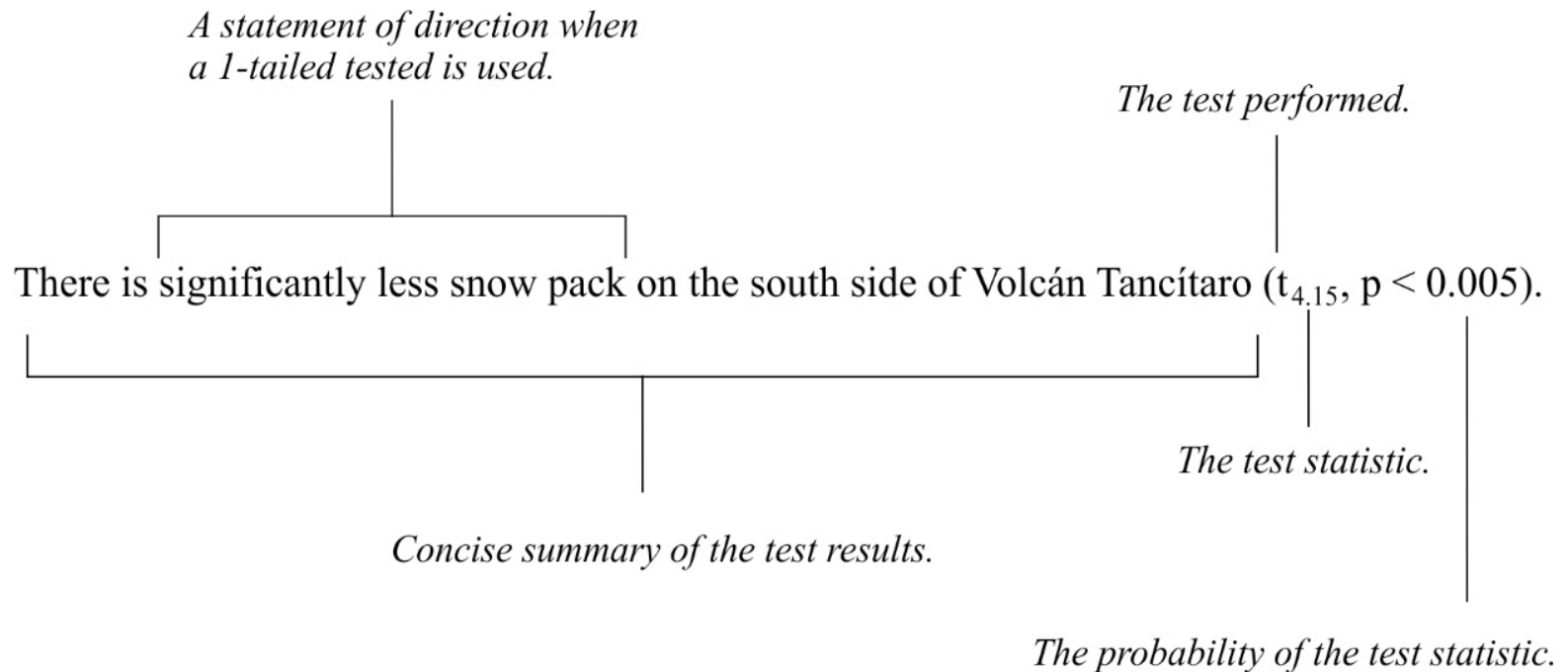
When a statistical test is performed there are two distinct, but related results:

- Test statistic: this is the value that is calculated.
- Probability: this is the probability of having that particular test statistic given the characteristics of the data set.

These characteristics are values such as the mean, standard deviation, sample size, etc...

Which characteristics are important depends on the specific test.

The results of any statistical test (e.g. one where you are testing a null hypothesis) must be stated in a concise summary statement. This statement should include a summary of the findings, the test that was performed, *the alpha level used IF different than 0.05*, the statistical results, and the probability.



Alpha Level (α)

The alpha level is the probability of committing a Type I error.

- A true null hypothesis that is incorrectly rejected.
- Also called the *significance* level.
- It is essentially a 'false positive'.

By convention we typically use 0.05 or 0.01 (5% or 1%) as our alpha level. *Therefore, we are willing to have 5 (or 1) out of 100 false positives.*

Beta Level (β)

The beta level is the probability of committing a Type II error.

- A false null hypothesis we fail to reject.
- It is essentially a 'false negative'.

By convention this value is not specified.

Type 1 and 2 Errors

	<i>If H_0 is true</i>	<i>If H_0 is false</i>
<i>And H_0 is rejected:</i>	Type I Error	No Error
<i>And H_0 is not rejected:</i>	No Error	Type II Error

Ho: There is no difference in survival rates between *Test Drug A* and a placebo.

In reality this is true... There really is no difference in survival.

A Type I error would occur if we incorrectly reject this true null hypothesis. *We are essentially saying that drug A works when it really doesn't.*

- The drug would go to market.
- People would take the drug expecting to survive.
- They would die since the drug has no effect.

This is bad...

Ho: There is no difference in survival rates between *Test Drug A* and a placebo.

In reality this is false... There really is difference in survival.

A Type II error would occur if we incorrectly accept this false null hypothesis. *We are essentially saying that drug A doesn't work when it actually does.*

- The drug would not go to market since our statistics showed it did not increase survival.
- People would never be given this drug.
- They would die since the drug has a positive effect but we said it did not.

This is also bad...

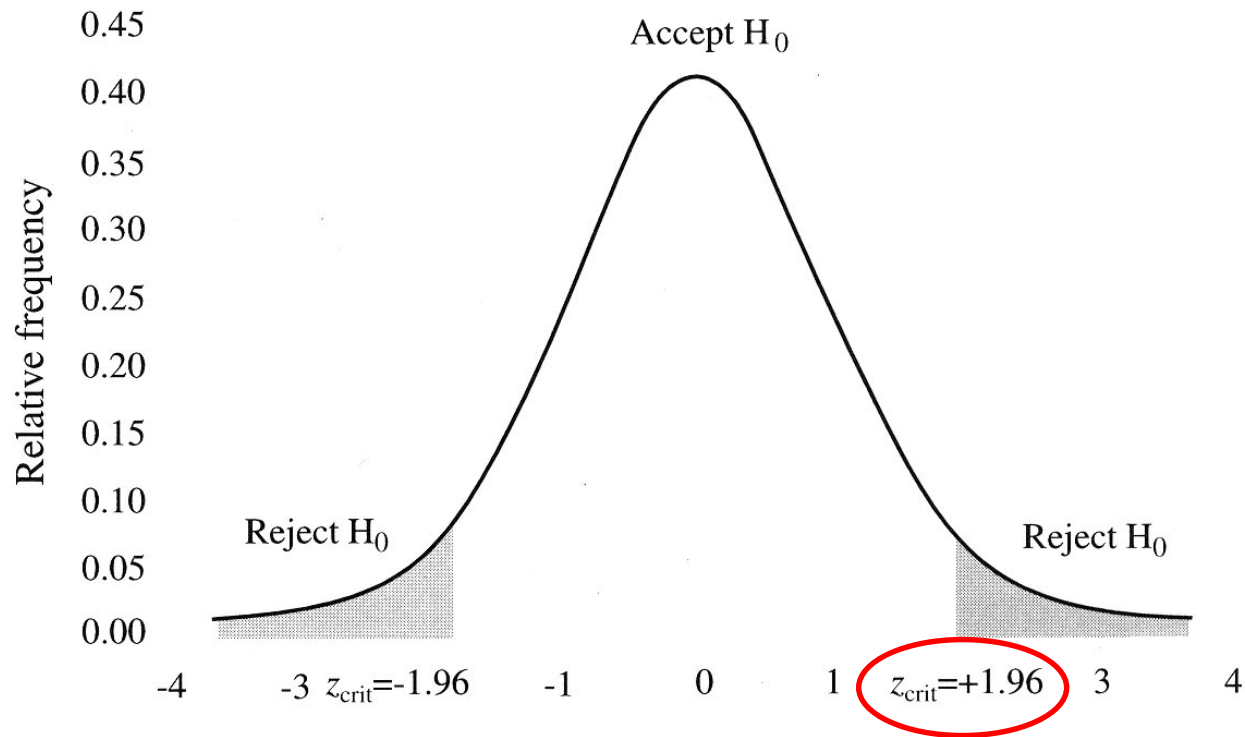
A Type I error accepts an alternate hypothesis when the results can be *attributed to chance*.

- So in effect we are stating that there is a difference when none actually exists.

A Type II error is only an error in the sense that we fail to correctly reject a false null hypothesis.

- It is not an error in the sense that an incorrect conclusion was drawn since no conclusion is drawn when the null hypothesis is accepted. *It is considered the lesser of two evils.*

Alpha level of 0.05



An α of 0.05 is equal to ± 1.96 sd from the mean.

One and Two-tailed Tests

- Two-tailed statistics test for difference *only*, direction must be *inferred*.

Ho: The rate of erosion at location A *is not* significantly different than the rate of erosion at location B.

Ha: The rate of erosion at location A *is* significantly different than the rate of erosion at location B.

- One-tailed statistics test for difference *and* '*a priori*' (prior knowledge) of the direction.

Ho: The rate of erosion at location A *is not* significantly *greater* than the rate of erosion at location B.

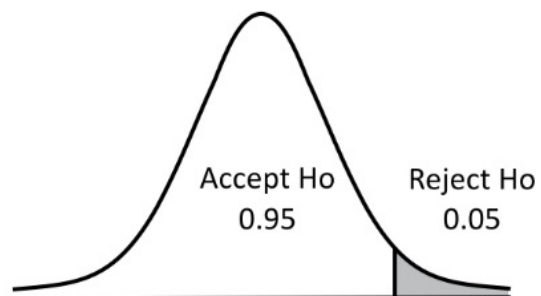
Ha: The rate of erosion at location A *is* significantly *greater* than the rate of erosion at location B.

Two-Tailed Test



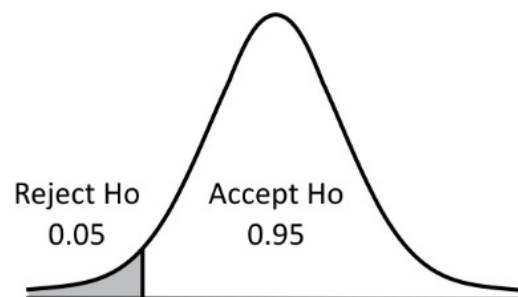
Difference can be negative or positive

One-Tailed Test



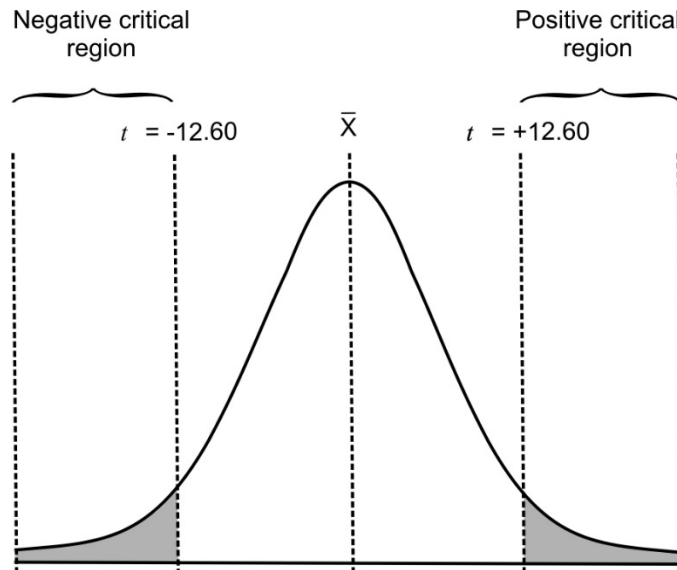
Difference is positive

One-Tailed Test

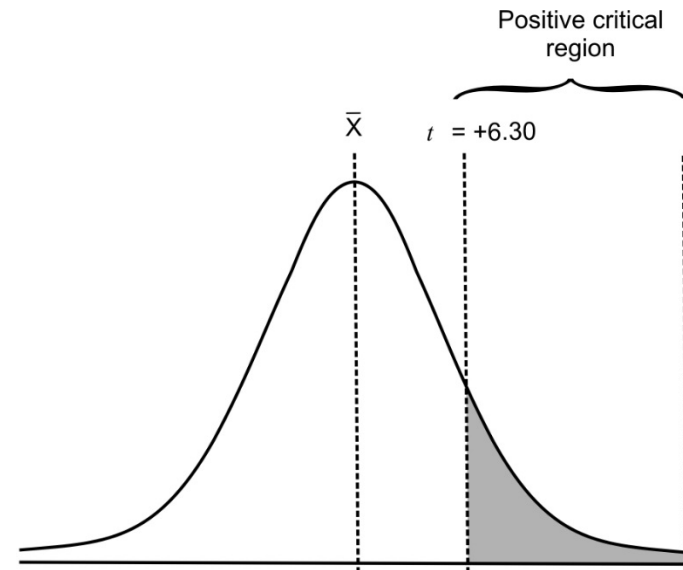


Difference is negative

Two-tailed test have the critical region split between positive and negative sides of the distribution. One-tailed tests do not. This is reflected in the table values (larger for two-tailed, smaller for one-tailed).



Two-Tailed Test



One-Tailed Test

Earlier research suggested that the average house value in York was lower than that in Lancaster.

Data for the average house value for each block group in downtown York and Lancaster were gathered from the census.

A two-sample t test was performed to determine whether housing values were lower in York than in Lancaster.

Since we have *a priori* (prior) knowledge of the direction of the difference (e.g. housing values are *less* in York) we would use a *one-tailed test*.

H_0 : Housing values in York are not significantly less than those in Lancaster.

H_a : Housing values in York are significantly less than those in Lancaster.

Note how the direction of the difference is stated in the hypotheses.

STOP!

Think about what was just stated: ***Housing values in York are not significantly less than those in Lancaster.***

You are 'expecting' that York's housing values may actually BE less, even though you are testing that they will not be less (remember we test the negative).

Question: What happens if you run the test and H_0 is correct: housing values are NOT less in York than Lancaster? *What if York's housing values are significantly GREATER Lancaster's... just the opposite of what you were expecting?*

You need to be prepared to explain why York's values were actually higher when you THOUGHT they'd be lower.

The test we will use is called a T-Test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} \quad \text{where} \quad s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} \quad s_p^2 = \frac{SS_1 + SS_2}{v_1 + v_2}$$

Housing Value (\$)

York	Lancaster
25368	49465
37045	37500
47500	53055
26785	48125
41493	45000
32864	52946
26140	

$$\alpha = 0.05$$

$$n_1 = 7 \quad n_2 = 6$$

$$df = (n_1 + n_2 - 2) = (7 + 6 - 2) = 11$$

$$\bar{X}_{York} = \frac{25368 + 37045 + 47500 + 26785 + 41493 + 32864 + 26140}{7} = 33885$$

$$\bar{X}_{Lancaster} = \frac{49465 + 37500 + 53055 + 48125 + 45000 + 52946}{6} = 47682$$

$$SS_{York} = (25368 - 33885)^2 + (37045 - 33885)^2 + (47500 - 33885)^2 + (26785 - 33885)^2 + (41493 - 33885)^2 + (32846 - 33885)^2 + (26140 - 33885)^2 = 437212244$$

$$SS_{Lancaster} = (49465 - 47682)^2 + (37500 - 47682)^2 + (53055 - 47682)^2 + (48125 - 47682)^2 + (45000 - 47682)^2 + (52946 - 47682)^2 = 2444393535$$

$$s_p^2 = \frac{437212244 + 2444393535}{6 + 5} = \frac{2881605779}{11} = 261964161.7$$

$$s_{X_1 - X_2} = \sqrt{\frac{261964161.7}{7} + \frac{261964161.7}{6}} = \sqrt{37423451.7 + 43660693.6} = \sqrt{81084145.3} = 9004.7$$

$$t = \frac{33885 - 47682}{9004.7} = -1.532 \quad (\text{ignore the sign when using the table})$$

$$t_{Critical} = 1.796 \quad \text{Since } 1.532 < 1.796, \text{ accept } H_0$$

Critical Values of the *t* Distribution

Taken from Zar, 1984 Table B.3

Tails



ν	$\alpha(2): 0.50$ $\alpha(1): 0.25$	0.20 0.10	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005	0.005 0.0025	0.002 0.001	0.001 0.0005
1	1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
31	0.682	1.309	1.696	2.040	2.453	2.744	3.022	3.375	3.633
32	0.682	1.309	1.694	2.037	2.449	2.738	3.015	3.365	3.622
33	0.682	1.308	1.692	2.035	2.445	2.733	3.008	3.356	3.611
34	0.682	1.307	1.691	2.032	2.441	2.728	3.002	3.348	3.601
35	0.682	1.306	1.690	2.030	2.438	2.724	2.996	3.340	3.591
36	0.681	1.306	1.688	2.028	2.434	2.719	2.990	3.333	3.582
37	0.681	1.305	1.687	2.026	2.431	2.715	2.985	3.326	3.574
38	0.681	1.304	1.686	2.024	2.429	2.712	2.980	3.319	3.566
39	0.681	1.304	1.685	2.023	2.426	2.708	2.976	3.313	3.558
40	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
41	0.681	1.303	1.683	2.020	2.421	2.701	2.967	3.301	3.544
42	0.680	1.302	1.682	2.018	2.418	2.698	2.963	3.296	3.538
43	0.680	1.302	1.681	2.017	2.416	2.695	2.959	3.291	3.532
44	0.680	1.301	1.680	2.015	2.414	2.692	2.956	3.286	3.526
45	0.680	1.301	1.679	2.014	2.412	2.690	2.952	3.281	3.520
46	0.680	1.300	1.679	2.013	2.410	2.687	2.949	3.277	3.515
47	0.680	1.300	1.678	2.012	2.408	2.685	2.946	3.273	3.510
48	0.680	1.299	1.677	2.011	2.407	2.682	2.943	3.269	3.505
49	0.680	1.299	1.677	2.010	2.405	2.680	2.940	3.265	3.500
50	0.679	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496

ν	$\alpha(2): 0.50$ $\alpha(1): 0.25$	0.20 0.10	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005	0.005 0.0025	0.002 0.001	0.001 0.0005
52	0.679	1.298	1.675	2.007	2.400	2.674	2.932	3.255	3.488
54	0.679	1.297	1.674	2.005	2.397	2.670	2.927	3.248	3.480
56	0.679	1.297	1.673	2.003	2.395	2.667	2.923	3.242	3.473
58	0.679	1.296	1.672	2.002	2.392	2.663	2.918	3.237	3.466
60	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
62	0.678	1.295	1.670	1.999	2.388	2.657	2.911	3.227	3.454
64	0.678	1.295	1.669	1.998	2.386	2.655	2.908	3.223	3.449
66	0.678	1.295	1.668	1.997	2.384	2.652	2.904	3.218	3.444
68	0.678	1.294	1.668	1.995	2.382	2.650	2.902	3.214	3.439
70	0.678	1.294	1.667	1.994	2.381	2.648	2.899	3.211	3.435
72	0.678	1.293	1.666	1.993	2.379	2.646	2.896	3.207	3.431
74	0.678	1.293	1.666	1.993	2.378	2.644	2.894	3.204	3.427
76	0.678	1.293	1.665	1.992	2.376	2.642	2.891	3.201	3.423
78	0.678	1.292	1.665	1.991	2.375	2.640	2.889	3.198	3.420
80	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
82	0.677	1.292	1.664	1.989	2.373	2.637	2.885	3.193	3.413
84	0.677	1.292	1.663	1.989	2.372	2.636	2.883	3.190	3.410
86	0.677	1.291	1.663	1.988	2.370	2.634	2.881	3.188	3.407
88	0.677	1.291	1.662	1.987	2.369	2.633	2.880	3.185	3.405
90	0.677	1.291	1.662	1.987	2.368	2.632	2.878	3.183	3.402
92	0.677	1.291	1.662	1.986	2.368	2.630	2.876	3.181	3.399
94	0.677	1.291	1.661	1.986	2.367	2.629	2.875	3.179	3.397
96	0.677	1.290	1.661	1.985	2.366	2.628	2.873	3.177	3.395
98	0.677	1.290	1.661	1.984	2.365	2.627	2.872	3.175	3.393
100	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
105	0.677	1.290	1.659	1.983	2.362	2.623	2.868	3.170	3.386
110	0.677	1.289	1.659	1.982	2.361	2.621	2.865	3.166	3.381
115	0.677	1.289	1.658	1.981	2.359	2.619	2.862	3.163	3.377
120	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
125	0.676	1.288	1.657	1.979	2.357	2.616	2.858	3.157	3.370
130	0.676	1.288	1.657	1.978	2.355	2.614	2.856	3.154	3.367
135	0.676	1.288	1.656	1.978	2.354	2.613	2.854	3.152	3.364
140	0.676	1.288	1.656	1.977	2.353	2.611	2.852	3.149	3.361
145	0.676	1.287	1.655	1.976	2.352	2.610	2.851	3.147	3.359
150	0.676	1.287	1.655	1.976	2.351	2.609	2.849	3.145	3.357
160	0.676	1.287	1.654	1.975	2.350	2.607	2.846	3.142	3.352
170	0.676	1.287	1.654	1.974	2.348	2.605	2.844	3.139	3.349
180	0.676	1.286	1.653	1.973	2.347	2.603	2.842	3.136	3.345
190	0.676	1.286	1.653	1.973	2.346	2.602	2.840	3.134	3.342
200	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
250	0.675	1.285	1.651	1.969	2.341	2.596	2.832	3.123	3.330
300	0.675	1.284	1.650	1.968	2.339	2.592	2.828	3.118	3.323
350	0.675	1.284	1.649	1.967	2.337	2.590	2.825	3.114	3.319
400	0.675	1.284	1.649	1.966	2.336	2.588	2.823	3.111	3.315
450	0.675	1.283	1.648	1.965	2.335	2.587	2.821	3.108	3.312
500	0.675	1.283	1.648	1.965	2.334	2.586	2.820	3.107	3.310
600	0.675	1.283	1.647	1.964	2.333	2.584	2.817	3.104	3.307
700	0.675	1.283	1.647	1.963	2.332	2.583	2.816	3.102	3.304
800	0.675	1.283	1.647	1.963	2.331	2.582	2.815	3.100	3.303
900	0.675	1.282	1.647	1.963	2.330	2.581	2.814	3.099	3.301
1000	0.675	1.282	1.646	1.962	2.330	2.581	2.813	3.098	3.300
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.0902	3.2905

$$t_{\text{Critical}} = 1.796$$

Since 1.532 < 1.796 we accept H_o .

Calculated t value.

Critical value from the table.

The calculated t value (1.532) is then compared to the critical t value (1.796).

- IF it is higher then we REJECT H_o .
- IF it is lower then we ACCEPT H_o .

The critical values are either taken from a table or calculated in SPSS.

Since in this class we will, by convention, use an alpha level of 0.05, our high power summary statement would read:

The housing values in York were not significantly less than the housing values in Lancaster ($t_{1.532}$, $0.10 > p > 0.05$).

Include the alpha level in the summary statement ONLY if it is different than 0.05. However, it must be stated somewhere.

The housing values in York were not significantly less than the housing values in Lancaster ($t_{1.532}, 0.10 > p > 0.05$).



How do we determine this?

This is called a *probability range*. If we are using a table it is rare that our calculated value will match the table values exactly.

The best we can do is state that our calculated value fell between two probabilities from table.

Tails



v	$\alpha(2):$	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
	$\alpha(1):$	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1		1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2		0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3		0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4		0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5		0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6		0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7		0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8		0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9		0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10		0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11		0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12		0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13		0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14		0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15		0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073

1.532 would fall about here on the table for a df of 11 at an alpha level of 0.05. So our probability range is between 0.10 and 0.05.

Probability Statements

1. Probability ranges are always written with the largest probability on the left, ***just as it appears in the table.***
2. $0.10 > p > 0.05$ means that our probability is less than 0.10, but greater than 0.05.
3. If you always place the largest probability to the left, then the greater than/less than symbols will always be $>$.
4. $0.10 < p < 0.05$ is nonsense... why?

If p is greater than 0.10, how can it ALSO be less than 0.05?

A few notes on tables:

1. Some tables are for both 1 and 2 tailed, others are not.
2. IF you are using a 1-tailed table for a 2-tailed test, you must split the alpha level between the tails:

2-tailed $\sigma = 0.05$... becomes $\sigma = 0.025$ on a 1-tailed table

3. Tables that are ONLY 2-tailed are typically for tests where direction can not be stated (e.g. AOV).

A few words on *degrees of freedom* (called *df* or *v*) ...

- Think of degrees of freedom as the minimum amount of information needed to be able to determine the value of ALL of the observations.
- For example, if we know the value of $n-1$ observations AND the mean, we can calculate the last observation value .

Data

5

$$n = 5$$

7

$$\bar{x} = 5$$

3

6

?

$$\frac{\sum x_i}{5} = 5$$

$$\sum x_i = 5 \times 5 = 25$$

$$25 - (5 + 7 + 3 + 6) = 4$$

In this example there is 1 degree of freedom.

Every time a statistical test is performed, you need to include the following:

1. A statement of the alpha level (somewhere).
2. Null and alternate hypotheses.
3. A high power summary statement that contains:
 - The test performed.
 - The calculated test statistic.
 - The exact probability (SPSS) or probability range (Table).

NOTE: Probabilities are *ALWAYS* written in descending order:

RIGHT: $0.10 > p > 0.05$

WRONG: ~~$0.05 < p < 0.10$~~

Developing Hypotheses Based on Maps

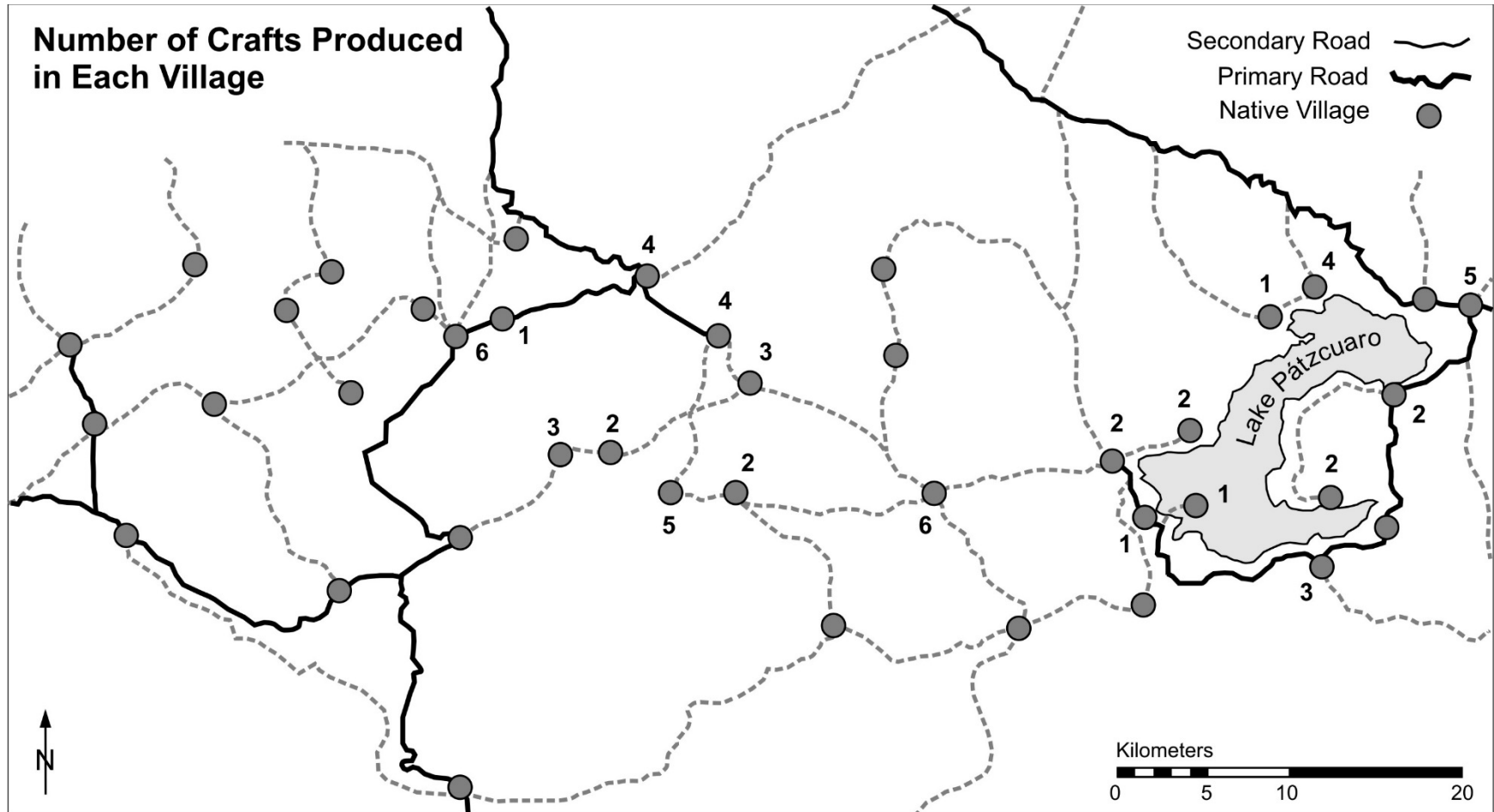
Maps can be used for:

- Delineating groups
- Locating observations
- Determining measurements

Measurements can be (but are not limited to):

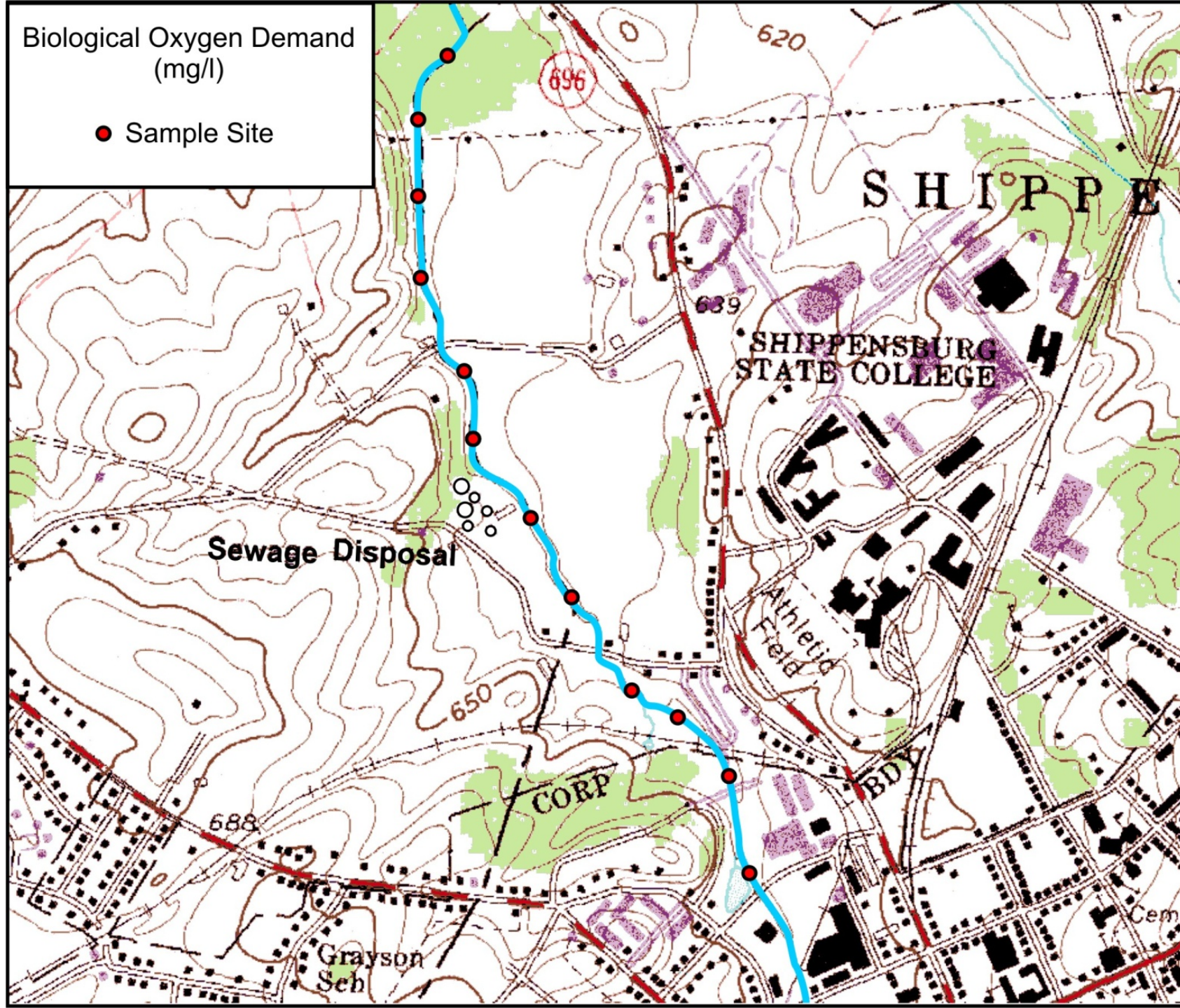
- Elevation
- Azimuth
- Aspect
- Distance
- Proximity

Develop null and alternate hypotheses based the following maps.

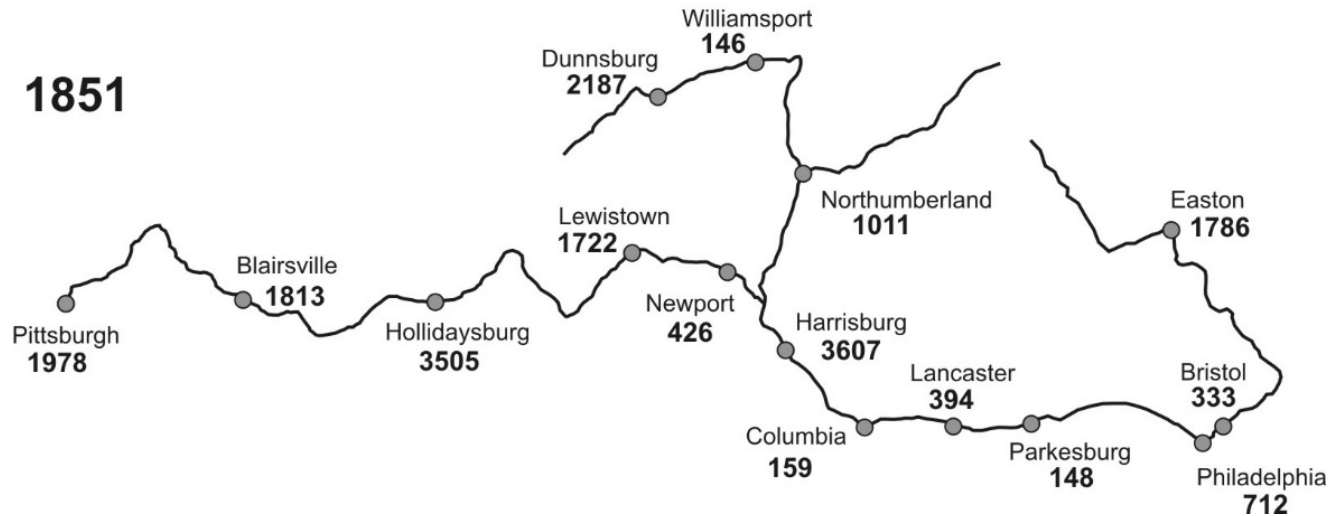
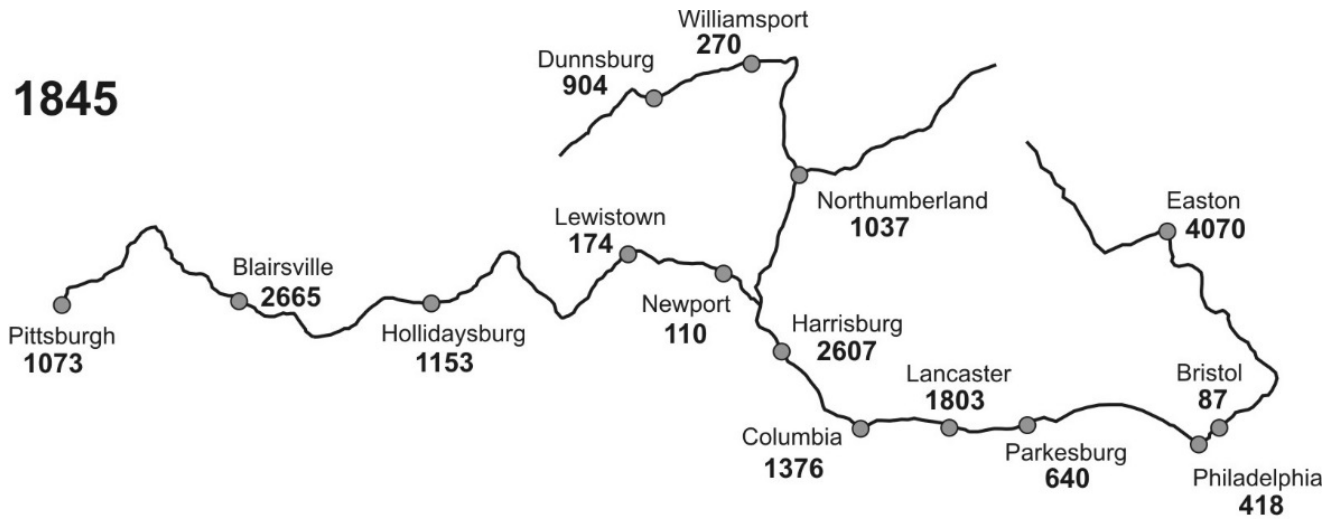


Biological Oxygen Demand
(mg/l)

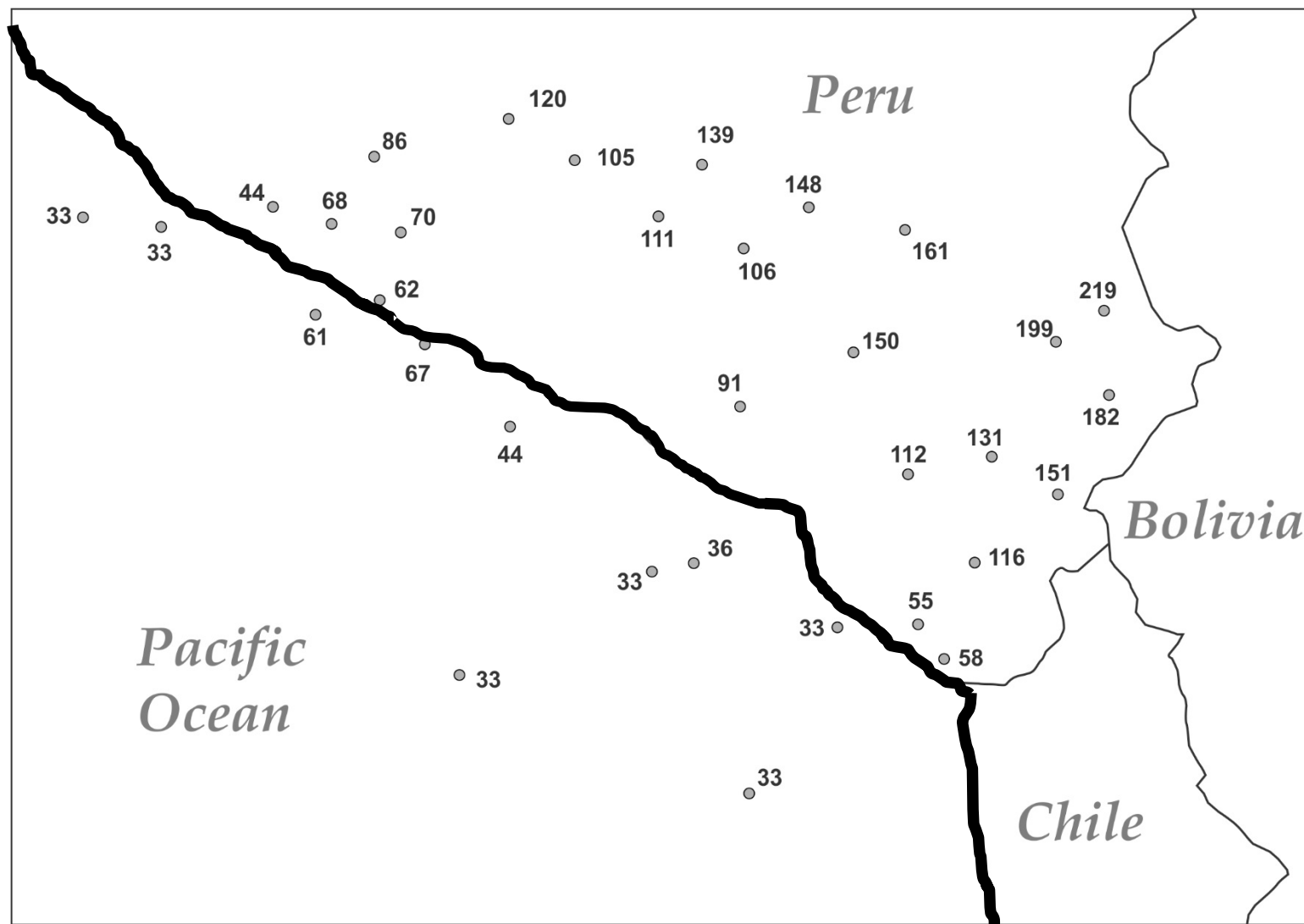
● Sample Site



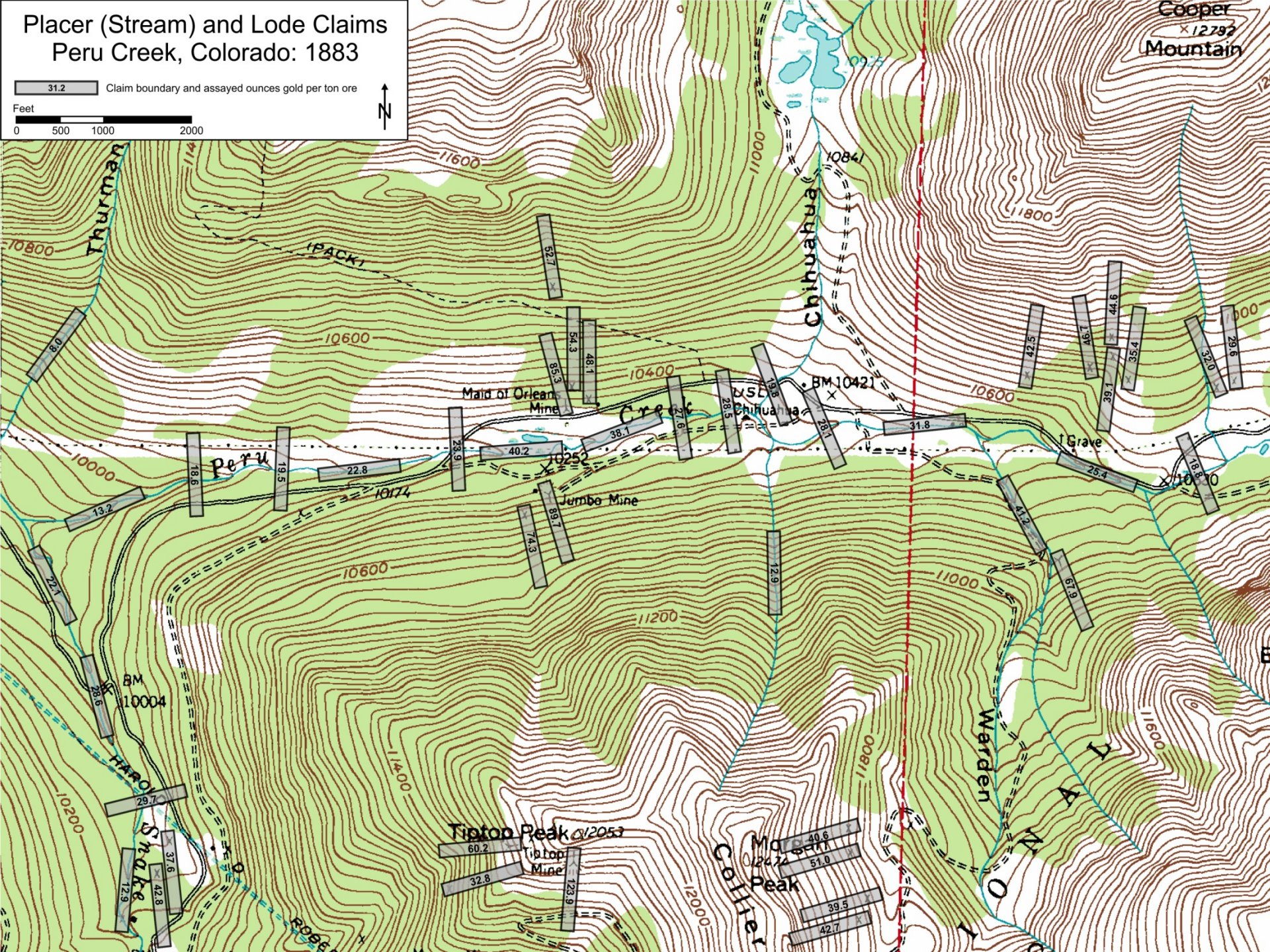
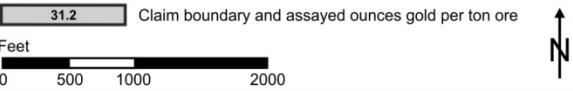
Ton of Pig Iron Shipped via Canal

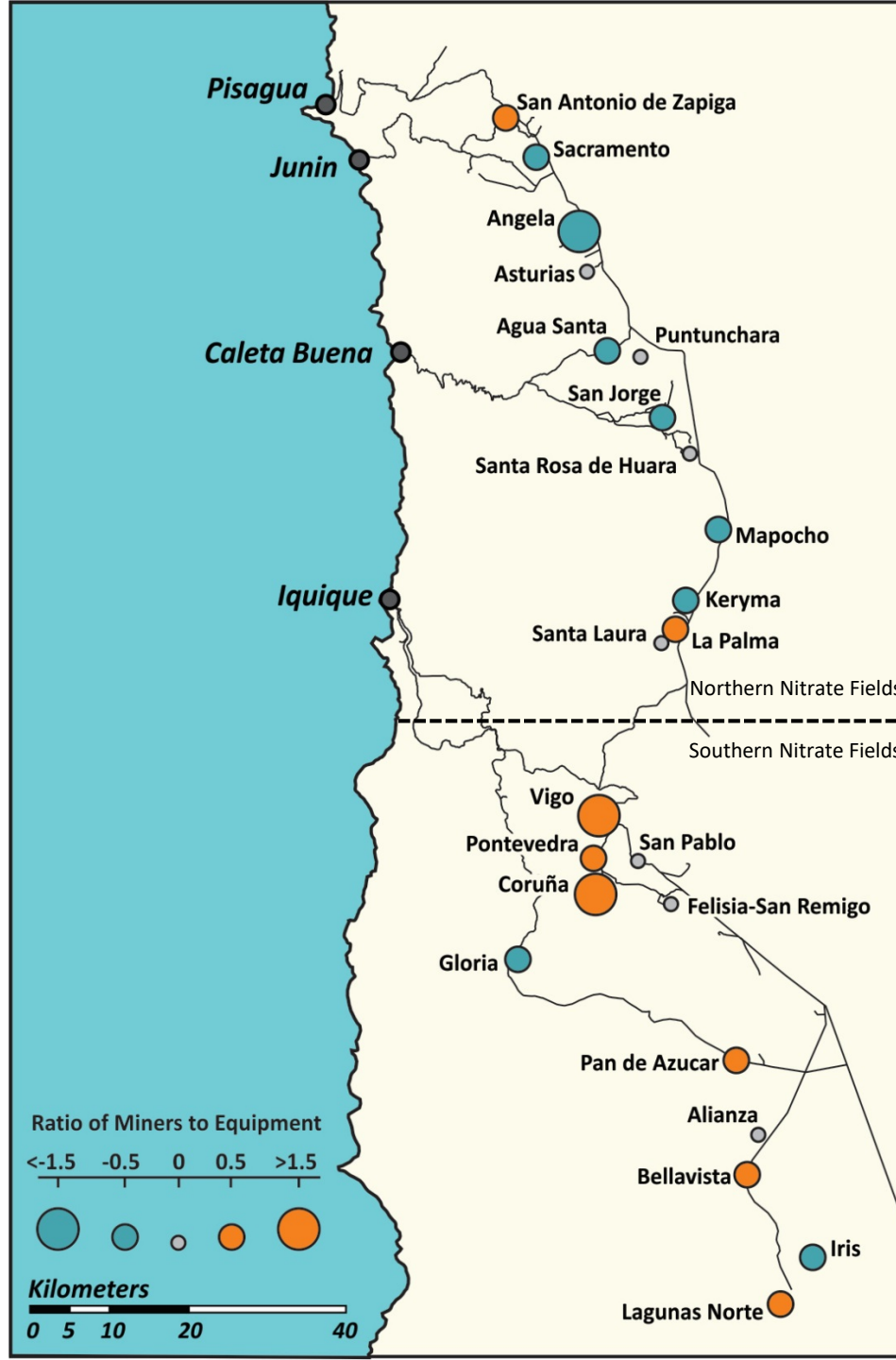


Earthquake Depths (km)



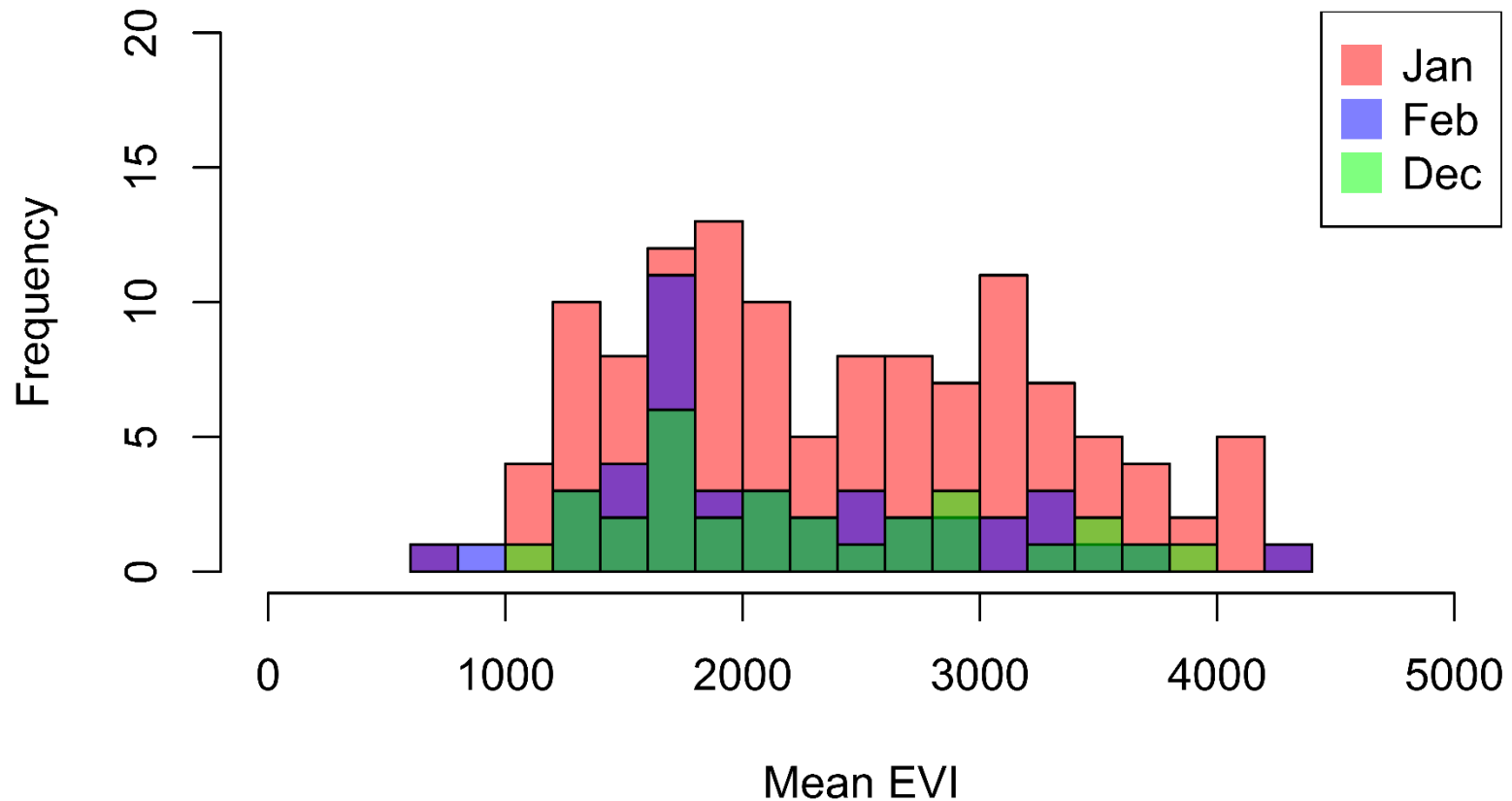
Placer (Stream) and Lode Claims Peru Creek, Colorado: 1883





What if there are 3 groups?

Mean EVI by Peak EVI Month



Often our hypotheses are not concerned with differences between or among groups.

- For example, an association between a measured variable and a measured landscape or natural characteristic.
- e.g. Decreasing temperature with increasing altitude.

In such cases, our null and alternate hypotheses might be:

H_o : *There no association between temperature and altitude.*

H_a : *There is an association between temperature and altitude.*

Data displayed as:
Housing prices per ft², distance
in meters to I-70

