

## Population/Sample Standard Deviation and Random Sampling

We selected Q3.2.20 (p.141) as an example of using StatCrunch to calculate population standard deviation and randomly select sample data from the population data then calculate sample standard deviation.

### Q3.2.20

**Travel Time** The following data represent the travel time (in minutes) to school for nine students enrolled in Sullivan's College Algebra course. Treat the nine students as a population.

Student	Travel Time	Student	Travel Time
Amanda	39	Scot	45
Amber	21	Erica	11
Tim	9	Tiffany	12
Mike	32	Glenn	39
Nicole	30		

- (a) Determine the population standard deviation.
- (b) Find three simple random sample of size 4 and determine the sample standard deviation of each sample.
- (c) Which samples underestimate the population standard deviation? Which overestimate the population standard deviation?

- (a) Determine the population standard deviation.

Step 1: Download the data set.

Step 2: Click **Stat** → **Summary Stats** → **Columns**.

The screenshot shows the StatCrunch web interface. The top navigation bar includes the Pearson logo, the StatCrunch logo, and the title 'Fundamentals of Statistics: Informed Decisions Using Data, Fourth Edition'. The main menu has buttons for StatCrunch, Applets, Edit, Data, Stat, Graph, and Help. The 'Stat' button is highlighted with a red box. A dropdown menu is open from 'Stat', showing options like Calculators, Summary Stats, Tables, Z Stats, T Stats, Proportion Stats, Variance Stats, Regression, ANOVA, Nonparametrics, Goodness-of-fit, Control Charts, and Resample. The 'Summary Stats' option is highlighted with a red box. A second dropdown menu is open from 'Summary Stats', showing options like Columns, Rows, Correlation, Covariance, and Grouped/Binned data. The 'Columns' option is highlighted with a red box. The background shows a data table with columns for Row, Student, and Travel Time, containing data for nine students.

- Step 3: 1) Click **Travel Time** under **Select Column(s)**:  
2) Choose **Unadj. Std. dev.** under **Statistics**:  
(In StatCrunch, Unadj. Std. dev. is the population standard deviation)  
3) Click **Compute!**

3\_2\_20.txt

Summary Stats

Select column(s):  
Travel Time

Where:  
Build

Group by:  
--optional--

Statistics:  
Unadj. variance  
Unadj. std. dev.  
Skewness  
Kurtosis  
Mode

Unadj. std. dev.

Percentiles (comma separated):  
Build

Other statistic (use x for data, e.g. mean(x)):  
Build

Output:  
☐ Store in data table

? Cancel Compute!

The population standard deviation is computed and shown below.

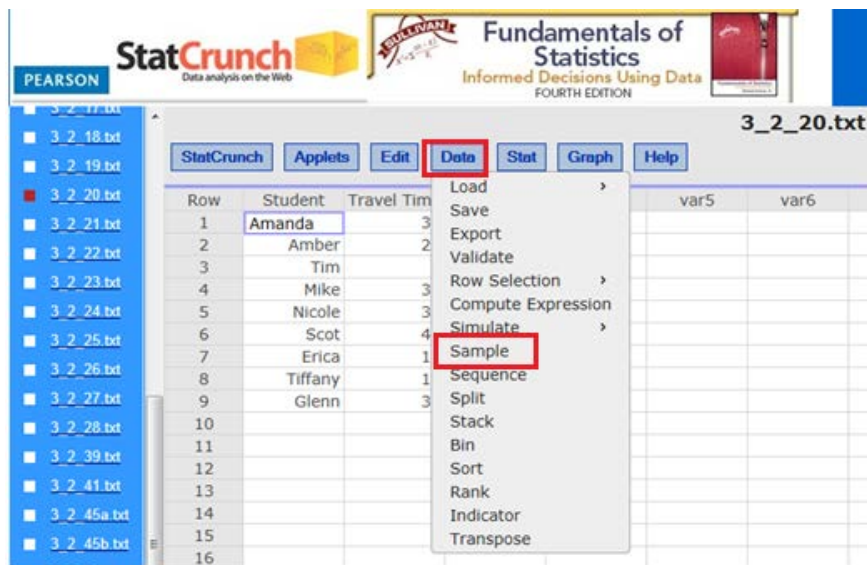
Options

Summary statistics:

Column	Unadj. std. dev.
Travel Time	12.841868

(b) Find three simple random sample of size 4 and determine the sample standard deviation of each sample.

Step 1: To find the three simple random sample of size 4, click **Data** → **Sample**.



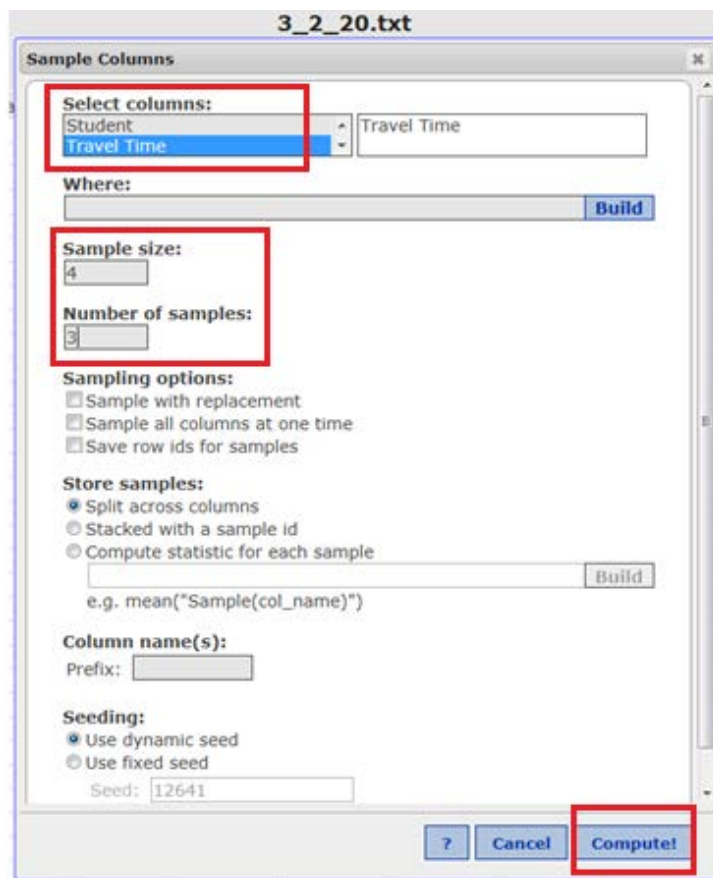
Step 2: 1) Click **Travel Time** under **Select Columns**:

2) Enter **4** for **Sample Size**: ---> **3** for **Number of Samples**:

(StatCrunch is randomly selecting the travel time of 4 people out of 9 people in the population and repeat the experiment 2 more times.)

3) Click **Compute!**

**Note:** Since the sample data are randomly selected, each time you perform step 2, you will obtain different sample data sets.



The three sample data sets are placed in three different columns.

3\_2\_20.txt

Row	Student	Travel Time	Sample1(Travel Time)	Sample2(Travel Time)	Sample3(Travel Time)
1	Amanda	39	30	21	39
2	Amber	21	39	39	45
3	Tim	9	45	32	39
4	Mike	32	12	12	30
5	Nicole	30			
6	Scot	45			
7	Erica	11			
8	Tiffany	12			
9	Glenn	39			
10					
11					

Step 3: To find sample standard deviation, click **Stat** → **Summary Stats** → **Columns**.

3\_2\_20.txt

Row	Student	Travel Time	Sample1(Travel Time)	Sample2(Travel Time)	Sample3(Travel Time)
1	Amanda	39	30	21	39
2	Amber	21	39	39	45
3	Tim	9	45	32	39
4	Mike	32	12	12	30
5	Nicole	30			
6	Scot	45			
7	Erica	11			
8	Tiffany	12			
9	Glenn	39			
10					
11					
12					
13					

Stat menu options:

- Calculators
- Summary Stats
- Tables
- Z Stats
- T Stats
- Proportion Stats
- Variance Stats
- Regression
- ANOVA
- Nonparametrics
- Goodness-of-fit
- Control Charts
- Resample

Summary Stats sub-menu options:

- Columns
- Rows
- Correlation
- Covariance
- Grouped/Binned data

- Step 4: 1) Under **Select column(s):**, choose **Sample1(Travel Time)**, **Sample2(Travel Time)**, and **Sample3(Travel Time)**. (Click while holding the **Ctrl** key on the keyboard)
- 2) Choose **Std. dev.** under **Statistics:** (Use Std. dev. for sample standard deviation)
- 3) Click **Compute!**

3\_2\_20.txt

**Summary Stats**

**Select column(s):**

- Travel Time
- Sample1(Travel Time)
- Sample2(Travel Time)
- Sample3(Travel Time)

**Where:**

**Group by:**

--optional--

**Statistics:**

- n
- Mean
- Variance
- Std. dev.
- Std. err.

**Percentiles (comma separated):**

**Other statistic (use x for data, e.g. mean(x)):**

**Output:**

☐ Store in data table

**Buttons:** ? Cancel **Compute!**

The three sample standard deviations are computed and shown below.

**Options**

**Summary statistics:**

Column	Std. dev.
Sample1(Travel Time)	14.387495
Sample2(Travel Time)	11.916375
Sample3(Travel Time)	6.1846584

(c) Which samples underestimate the population standard deviation? Which overestimate the population standard deviation?

From part (b), sample standard deviation for Sample 1 is 14.39, for Sample 2 is 11.92, and for Sample 3 is 6.18. The population standard deviation, from part (a), is 12.84. Therefore, Sample 1 overestimates and Sample 2 and 3 underestimate the population standard deviation.