

Assignment 8

March 7, 2021

0.0.1 Portfolio assignment 8

15 min: - Calculate the 90%, 95%, 99% and 99.99% confidence interval for at least 2 columns with numerical data in the dataset that you chose in portfolio assignment 4. Do you see the impact the confidence has on the interval? - Now calculate the 95% confidence interval again but use only the first 10% of your rows. Compare this interval to the previous 95% confidence interval you calculated. Do you see the impact of having less data?

```
[1]: import pandas as pd
import numpy as np
import scipy.stats as st
```

```
[2]: steam = pd.read_csv("steam_games.csv")
steam.dropna(inplace=True) # we remove all Na/NaN values
steam.head()
```

```
[2]:
```

	url	types	\
6	https://store.steampowered.com/app/601150/Devi...	app	
37	https://store.steampowered.com/app/532210/Life...	app	
75	https://store.steampowered.com/app/476600/Call...	app	
94	https://store.steampowered.com/app/418240/Shad...	app	
153	https://store.steampowered.com/app/513710/SCUM/	app	

	name	\
6	Devil May Cry 5	
37	Life is Strange 2	
75	Call of Duty®: WWII	
94	Shadow Tactics: Blades of the Shogun	
153	SCUM	

	desc_snippet	\
6	The ultimate Devil Hunter is back in style, in...	
37	After a tragic incident, brothers Sean and Dan...	
75	Call of Duty® returns to its roots with Call o...	
94	Shadow Tactics is a hardcore tactical stealth ...	
153	SCUM aims to evolve the multiplayer open world...	

	recent_reviews	\
--	----------------	---

6 Very Positive,(408),- 87% of the 408 user revi...
 37 Mostly Positive,(155),- 72% of the 155 user re...
 75 Mixed,(66),- 53% of the 66 user reviews in the...
 94 Very Positive,(124),- 94% of the 124 user revi...
 153 Mixed,(346),- 58% of the 346 user reviews in t...

all_reviews release_date \
 6 Very Positive,(9,645),- 92% of the 9,645 user ... Mar 7, 2019
 37 Very Positive,(4,437),*- 86% of the 4,437 use... Sep 26, 2018
 75 Mixed,(12,703),- 54% of the 12,703 user review... Nov 2, 2017
 94 Overwhelmingly Positive,(7,238),- 95% of the 7... Dec 6, 2016
 153 Mixed,(22,177),- 64% of the 22,177 user review... Aug 29, 2018

developer \
 6 CAPCOM Co., Ltd.
 37 DONTNOD Entertainment
 75 Sledgehammer Games,Raven Software
 94 Mimimi Productions
 153 Gamepires,Croteam

publisher \
 6 CAPCOM Co., Ltd.,CAPCOM Co., Ltd.
 37 Square Enix,Square Enix
 75 Activision,Activision
 94 Daedalic Entertainment,Daedalic Entertainment
 153 Devolver Digital,Devolver Digital

popular_tags \
 6 Action,Hack and Slash,Great Soundtrack,Demons,...
 37 Choices Matter,Story Rich,Adventure,Great Soun...
 75 World War II,FPS,Multiplayer,War,Action,Zombie...
 94 Stealth,Tactical,Ninja,Strategy,Real Time Tact...
 153 Early Access,Survival,Open World,Multiplayer,Z...

game_details \
 6 Single-player,Online Multi-Player,Online Co-op...
 37 Single-player,Steam Achievements,Full controll...
 75 Single-player,Multi-player,Online Multi-Player...
 94 Single-player,Steam Achievements,Full controll...
 153 Single-player,Multi-player,Online Multi-Player...

languages achievements \
 6 English,French,Italian,German,Spanish - Spain,... 51.0
 37 English,French,Italian,German,Spanish - Spain,... 46.0
 75 English,French,Italian,German,Spanish - Spain,... 35.0
 94 English,French,German,Spanish - Spain,Japanese... 44.0
 153 English,German,Russian,Simplified Chinese,Fren... 56.0

		genre \
6		Action
37		Adventure
75		Action
94		Indie,Strategy
153	Action,Adventure,Indie,Massively Multiplayer,E...	

		game_description \
6	About This Game The Devil you know returns in...	
37	About This Game Purchase the Complete Season...	
75	About This Game Includes Digital Edition Bon...	
94	About This Game Shadow Tactics is a hardcore...	
153	About This Game The world's unquenchable need...	

		mature_content \
6	Mature Content Description The developers de...	
37	Mature Content Description The developers de...	
75	Mature Content Description The developers de...	
94	Mature Content Description The developers de...	
153	Mature Content Description The developers de...	

		minimum_requirements \
6	Minimum:,OS:,WINDOWS® 7, 8.1, 10 (64-BIT Requi...	
37	Minimum:,OS:,Windows 7 or above (64-bit Operat...	
75	Minimum:,Requires a 64-bit processor and opera...	
94	Minimum:,OS:,Windows Vista 64-bit,Processor:,I...	
153	Minimum:,Requires a 64-bit processor and opera...	

		recommended_requirements	original_price \
6	Recommended:,OS:,WINDOWS® 7, 8.1, 10 (64-BIT R...		\$59.99
37	Recommended:,OS:,Windows 10 64-bit,Processor:,...		\$7.99
75	Recommended:,Requires a 64-bit processor and o...		1.020
94	Recommended:,OS:,Windows 7 64-bit or higher, W...		\$39.99
153	Recommended:,Requires a 64-bit processor and o...		\$19.99

	discount_price
6	\$70.42
37	\$39.95
75	\$906.48
94	\$13.59
153	\$28.48

0.1 Test 1 - Achievements

```
[3]: steam['achievements'].mean()
```

[3]: 132.90243902439025

```
[4]: # confidence 90%  
st.t.interval(0.9, len(steam)-1, loc=steam['achievements'].mean(), scale=st.  
    ↪sem(steam['achievements']))
```

[4]: (41.63202451662629, 224.17285353215416)

```
[5]: # Confidence 95%  
st.t.interval(0.95, len(steam)-1, loc=steam['achievements'].mean(), scale=st.  
    ↪sem(steam['achievements']))
```

[5]: (23.76051268042771, 242.04436536835277)

```
[6]: # Confidence 0.99  
st.t.interval(0.99, len(steam)-1, loc=steam['achievements'].mean(), scale=st.  
    ↪sem(steam['achievements']))
```

[6]: (-11.796335262232219, 277.6012133110127)

```
[7]: # Confidence 99.99%  
st.t.interval(0.9999, len(steam)-1, loc=steam['achievements'].mean(), scale=st.  
    ↪sem(steam['achievements']))
```

[7]: (-91.62776936310345, 357.432647411884)

0.2 Test 2 - Discount

```
[8]: steam['discount_price'] = steam['discount_price'].str.replace('$', '').  
    ↪astype(float) # make type float
```

```
[9]: steam['discount_price'].mean()
```

[9]: 63.91243902439022

```
[10]: # confidence 90%  
st.t.interval(0.9, len(steam)-1, loc=steam['discount_price'].mean(), scale=st.  
    ↪sem(steam['discount_price']))
```

[10]: (43.524371102032006, 84.30050694674841)

```
[11]: # confidence 95%  
st.t.interval(0.95, len(steam)-1, loc=steam['discount_price'].mean(), scale=st.  
    ↪sem(steam['discount_price']))
```

[11]: (39.53221658891045, 88.29266145986998)

```
[12]: # confidence 99%
      st.t.interval(0.99, len(steam)-1, loc=steam['discount_price'].mean(), scale=st.
      ↪sem(steam['discount_price']))
```

```
[12]: (31.589495665930023, 96.23538238285042)
```

```
[13]: # confidence 90.99%
      st.t.interval(0.9999, len(steam)-1, loc=steam['discount_price'].mean(),
      ↪scale=st.sem(steam['discount_price']))
```

```
[13]: (13.756677555419827, 114.0682004933606)
```

Conclusion: The range between the interval seems to be increasing whenever we make our confidence higher

0.3 Test 3 - Limited data

```
[14]: limited = steam.head(int(len(steam)*(10/100)))
```

```
[15]: st.t.interval(0.95, len(steam)-1, loc=steam['achievements'].mean(), scale=st.
      ↪sem(steam['achievements']))
      # full
```

```
[15]: (23.76051268042771, 242.04436536835277)
```

```
[16]: st.t.interval(0.95, len(limited)-1, loc=limited['achievements'].mean(),
      ↪scale=st.sem(limited['achievements']))
      # limited
```

```
[16]: (36.98004338651619, 92.5199566134838)
```

Conclusion: limiting the dataset decreased the range