Football part 1

Ahmed Ayman

October 17, 2022

Agenda

1. The question
2. The accepted answer
3. Improve: compare best
4. Improve: Non parametric
5. The power of business understanding
6. Answers evaluation
7. Manger reaction

The question

Ahmed Ayman • You

Data man

2mo •

لو افترضنا ان الزمالك خد الدورى

وانا عندى سؤال
- هل الزمالك خد الدورى عشان الزمالك حلو؟
- ولا خد الدورى عشان الاهلى وحش؟

ازاى اقدر اجاوب ع السؤال ده بالData science؟

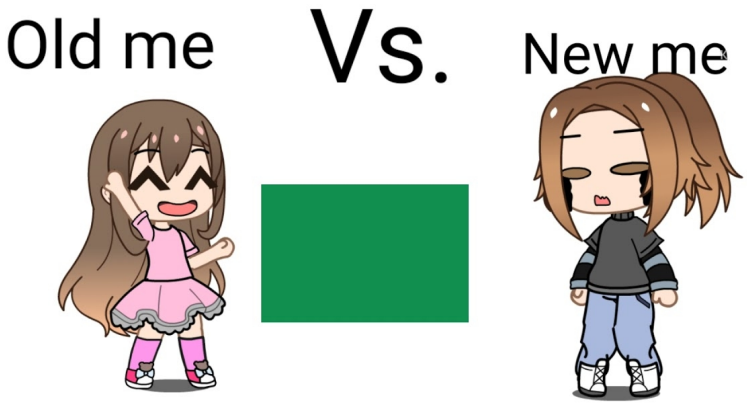
ياريت محدش يدخلنى فى نقاشات كرويه او تعصب كروى
انا حابب اشوف الافكار من وجهة نظر احصائيه 😅

The accepted answer

- compare Zamalek result with old years results
- compare Ahly result with old years results

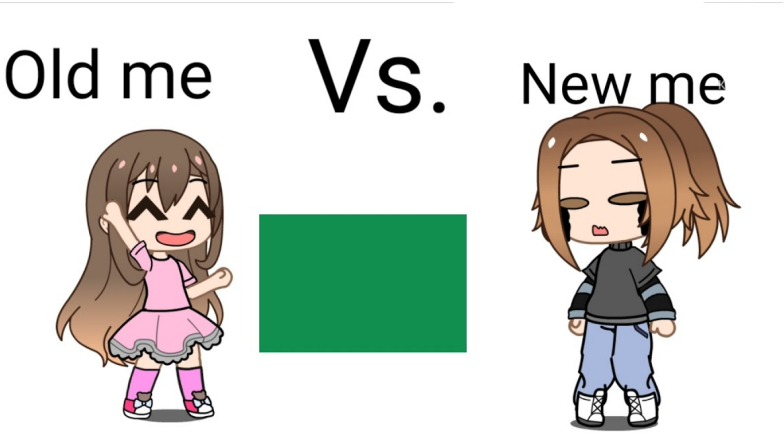
The accepted answer

- compare Zamalek result with old years results
- compare Ahly result with old years results



The accepted answer

- compare Zamalek result with old years results
- compare Ahly result with old years results



compare best

I get the data from the archive

season	p	play	p%
2022	77	34	**75%**
2021	80	34	**78%**
2020	89	34	**87%**
2019	80	34	**78%**
2018	88	34	**86%**
2017	84	34	**82%**
2016	76	34	**76%**
2015	87	38	**76%**
2011	61	30	**68%**
2010	65	30	**72%**
2009	63	30	**70%**
2008	70	30	**78%**
2007	73	30	**81%**
2006	72	26	**92%**
2005	74	26	**95%**

compare best

I get the data from the archive

season	p	play	p%
2022	77	34	**75%**
2021	80	34	**78%**
2020	89	34	**87%**
2019	80	34	**78%**
2018	88	34	**86%**
2017	84	34	**82%**
2016	76	34	**76%**
2015	87	38	**76%**
2011	61	30	**68%**
2010	65	30	**72%**
2009	63	30	**70%**
2008	70	30	**78%**
2007	73	30	**81%**
2006	72	26	**92%**
2005	74	26	**95%**

count	13
min	70%
mean	81%
median	84%
max	95%

I'll split into 3 clusters

Bad	(70 - 78.3)
Acceptable	(78.3 - 86.6)
Good	(86.6 - 94.9)

compare best

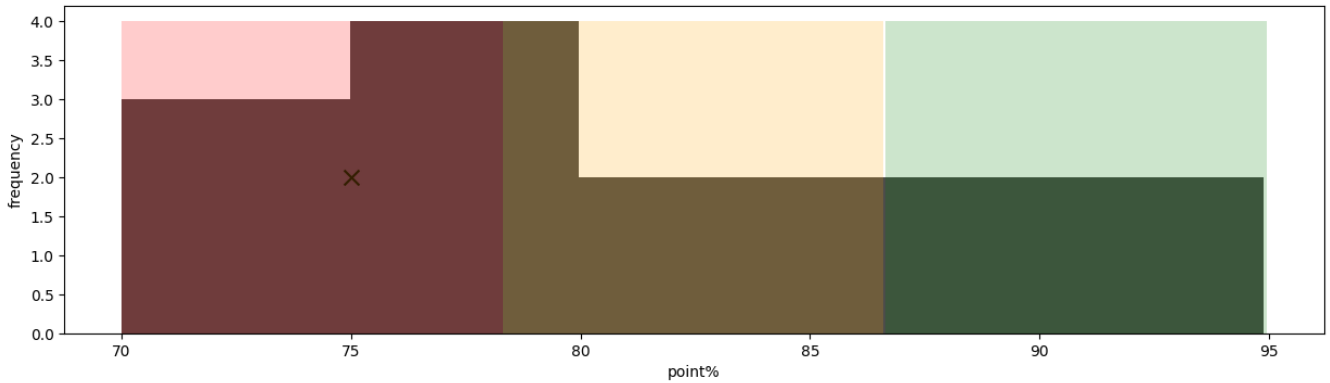
I get the data from the archive

season	p	play	p%
2022	77	34	**75%**
2021	80	34	**78%**
2020	89	34	**87%**
2019	80	34	**78%**
2018	88	34	**86%**
2017	84	34	**82%**
2016	76	34	**76%**
2015	87	38	**76%**
2011	61	30	**68%**
2010	65	30	**72%**
2009	63	30	**70%**
2008	70	30	**78%**
2007	73	30	**81%**
2006	72	26	**92%**
2005	74	26	**95%**

count	13
min	70%
mean	81%
median	84%
max	95%

I'll split into 3 clusters

Bad	(70 - 78.3)
Acceptable	(78.3 - 86.6)
Good	(86.6 - 94.9)



- Zamalek, not above expectation (under average)
- The main reason for wining, because Ahly was pathetic