DM & Visualization - Exam 2020-02-14 - Solution

Figure 1: Modified from source: https://www.idealmedicalcare.org/1103-doctor-suicides-13-reasons-why/

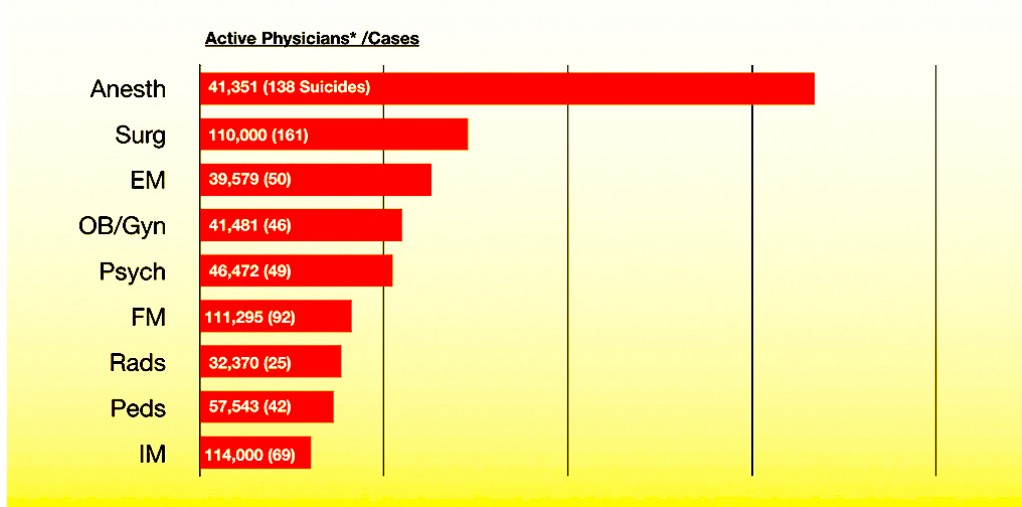## Analysis

Analyze the above graph that was published on a medical blog in 2018.

Question: Is there a clearly defined question addressed by the visualization? Write it down.

Yes: what the incidence of suicides among different medical specialties?

Data: Is the data quality appropriate? Circle the inadequate characteristics and explain.

Characteristic	Adequacy assessment
Accuracy	*Partly*: no. of Surg and IM are too round to be accurate
Completeness	Yes: we assume all specialties are reported
Consistency	**NO**: the sum of suicides is not 1103 as reported in the title
−	+: no. of physicians is from 2016 suicides are presumably on a wider time frame
Currency	*Partly*: Data is from 2018 (2016 the active)
Credibility	Yes: data seem to come from trustable sources
Understandability	**No**: The length of the bar encodes neither the suicide cases nor the no. of active physicians. One could think it is the ratio Active Physicians / Cases but actually it is the reciprocal of that ratio. The value is the suicide rate.
Precision	Yes Precision seems reasonable for the purpose

Visual Proportionality: Are the values encoded in a uniformly proportional way?

Apparently the representation is proportional.

Visual Utility: All the elements in the graph convey useful information?

- The gradient background is not useful.
- The strongly bright colors are not useful and may lead to sight fatigue.
- The vertical lines are not much useful without a proper axis.

Visual Clarity: Are the data in the graph clearly identifiable and understandable (properly described)?

The data reported in each bar might support the comprehension. Though we miss the value that is encoded in the bar lengths: it is not immediate to compute the ratios and compare them (e.g. 50 / 39579).

The note above the graph is clearly misleading because it suggest the reciprocal of the rate is used.

Design

Design the visualization based on the following data structure (to be completed)

| Field	Dim./Measure	Description
SPECIALITY	Dimension	Different medical specialties
ACTIVE_PHYSICIANS	Measure	No. of active physicians in that specialty
SUICIDES	Measure	No. of suicides among physicians in that specialty

Design schema

Fill in the required schema elements; formulas can be used if required.

Schema	Details
Columns:	SUM(SUICIDES) / SUM(ACTIVE_PHYSICIANS) * 10000
Rows:	SPECIALITY
Graph type:	Bar
Color:	
Size:	
Label:	SUM(SUICIDES) / SUM(ACTIVE_PHYSICIANS) * 10000

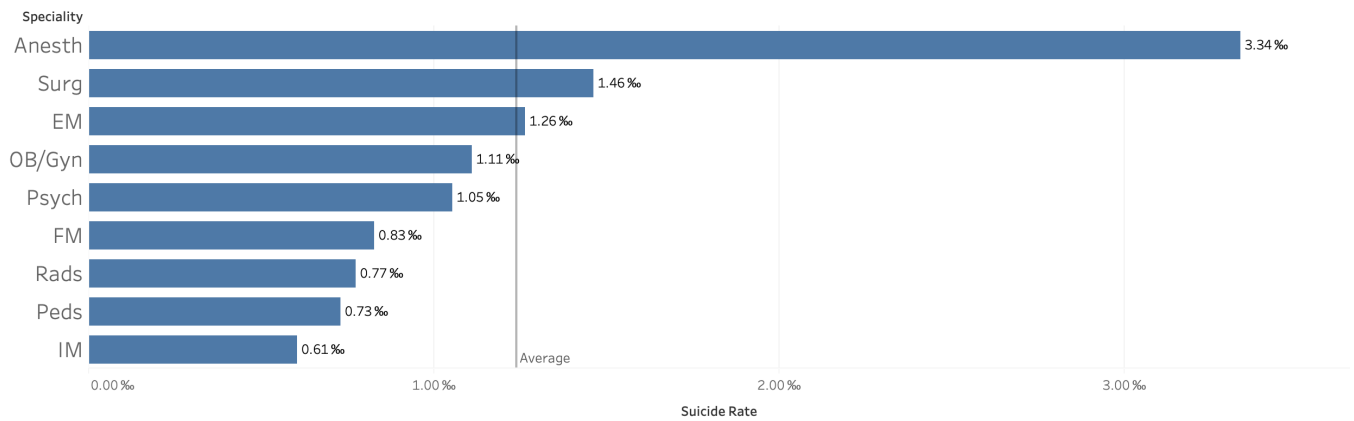**Sketch of the resulting graph:**

Bar - Rate PTT
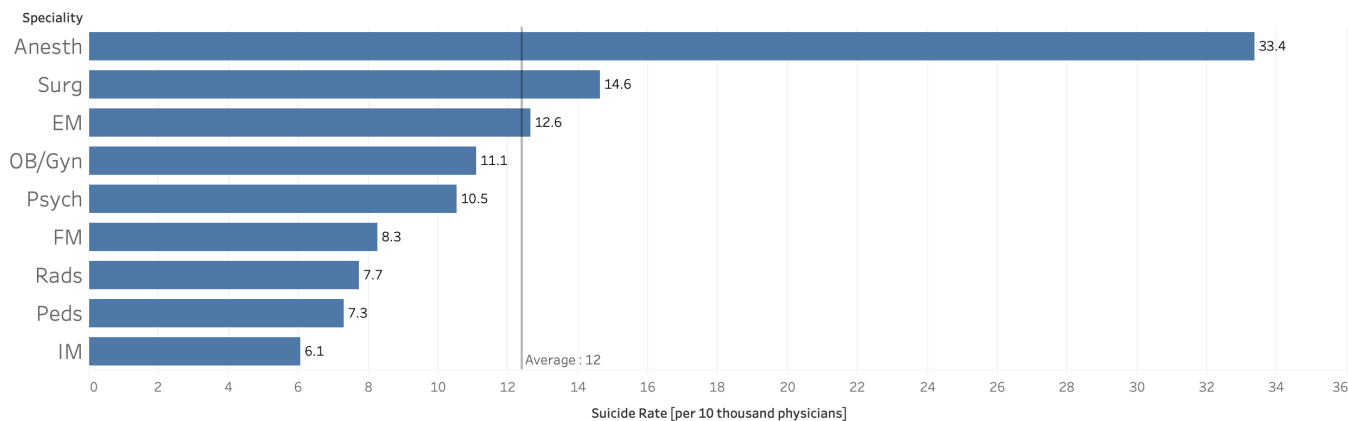
Sum of Suicide Rate [per 10 thousand physicians] for each Speciality. The marks are labeled by sum of Suicide Rate [per 10 thousand physicians].
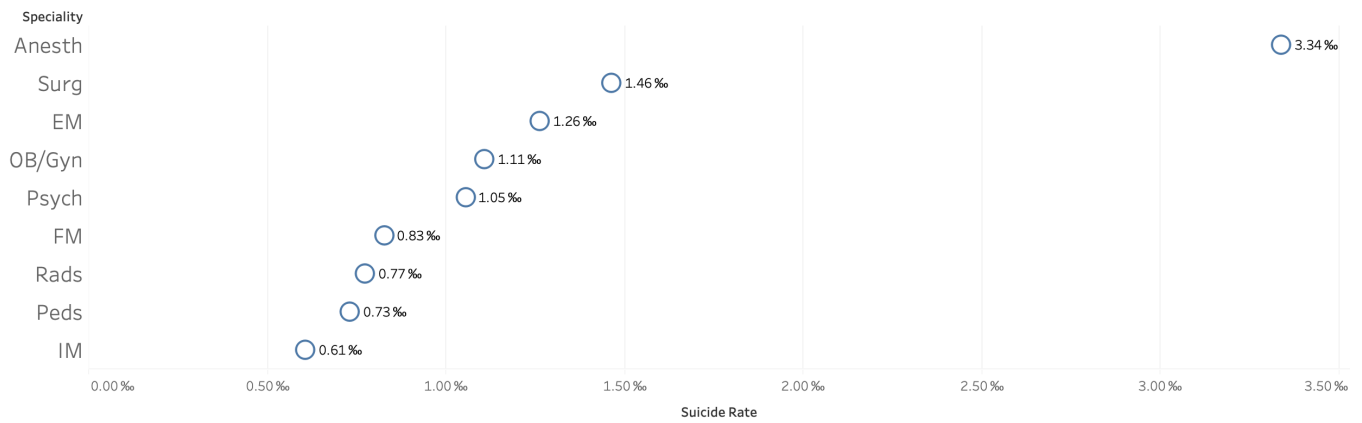
Other options:

Same as above using ‰

Bar - Rate

Speciality

	Suicide Rate
Anesth	3.34 ‰
Surg	1.46 ‰
EM	1.26 ‰
OB/Gyn	1.11 ‰
Psych	1.05 ‰
FM	0.83 ‰
Rads	0.77 ‰
Peds	0.73 ‰
IM	0.61 ‰

0.00 ‰   1.00 ‰   Average   2.00 ‰   3.00 ‰

Suicide Rate

Sum of Suicide Rate for each Speciality. The marks are labeled by sum of Suicide Rate.

Using a dot plot:

Dots - Rate

Speciality

	Suicide Rate
Anesth	3.34 ‰
Surg	1.46 ‰
EM	1.26 ‰
OB/Gyn	1.11 ‰
Psych	1.05 ‰
FM	0.83 ‰
Rads	0.77 ‰
Peds	0.73 ‰
IM	0.61 ‰

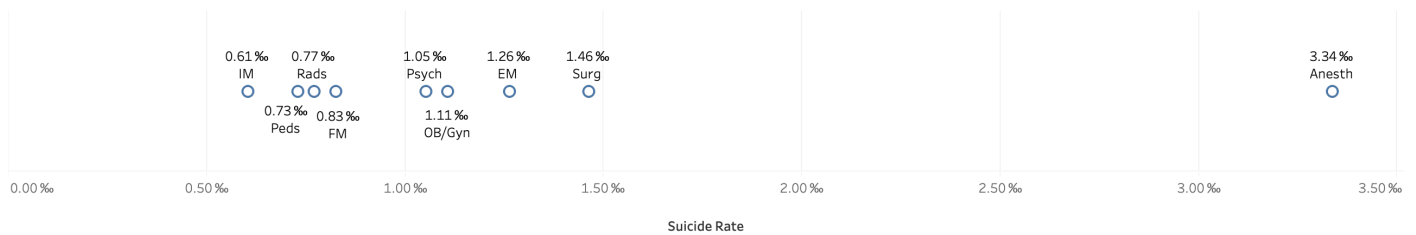0.00 ‰   0.50 ‰   1.00 ‰   1.50 ‰   2.00 ‰   2.50 ‰   3.00 ‰   3.50 ‰

Suicide Rate

Sum of Suicide Rate for each Speciality. The marks are labeled by sum of Suicide Rate.

Using a dot strip:

DotStrip - Rate

0.61 ‰ IM   0.77 ‰ Rads   1.05 ‰ Psych   1.26 ‰ EM   1.46 ‰ Surg   3.34 ‰ Anesth

0.73 ‰ Peds   0.83 ‰ FM   1.11 ‰ OB/Gyn

0.00 ‰   0.50 ‰   1.00 ‰   1.50 ‰   2.00 ‰   2.50 ‰   3.00 ‰   3.50 ‰

Suicide Rate

Sum of Suicide Rate. The marks are labeled by sum of Suicide Rate and Speciality.