

## Application exercise:

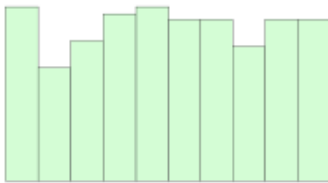
### Distributions of numerical variables

WRITE YOUR RESPONSES ON A PIECE OF PAPER. WRITE LEGIBLY! ONLY ONE SUBMISSION PER TEAM IS REQUIRED.

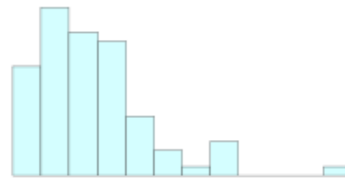
### Shapes of distributions

- Below are two histograms. One corresponds to the age at which a sample of people applied for marriage licenses; the other corresponds to the last digit of a sample of social security numbers. Which graph is which, and why?

(a)



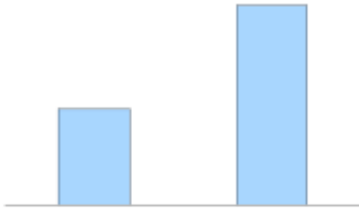
(b)



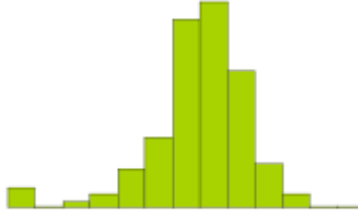
*(a) SSN, no pattern expected so uniform (b) age at first marriage, more people get married earlier*

- Match the following variables with the histograms and bar graphs given below. These data represent Sta 101 students at Duke. [*Hint: Think about how each variable should behave.*]
- |                                                                      |                                                               |
|----------------------------------------------------------------------|---------------------------------------------------------------|
| (a) the height of students                                           | (d) the number of hours of sleep students received last night |
| (b) gender breakdown of students                                     | (e) whether or not students live off campus                   |
| (c) the time it took students to get to their first class of the day | (f) the number of piercings students have                     |

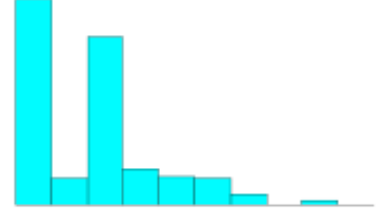
(1)



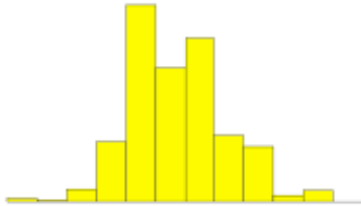
(3)



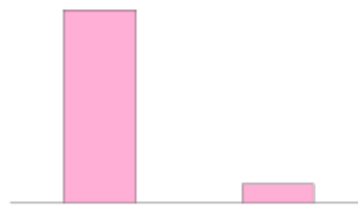
(5)



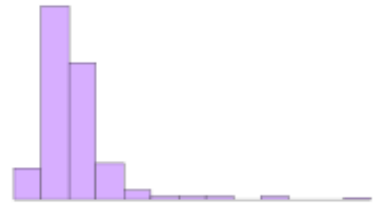
(2)



(4)



(6)

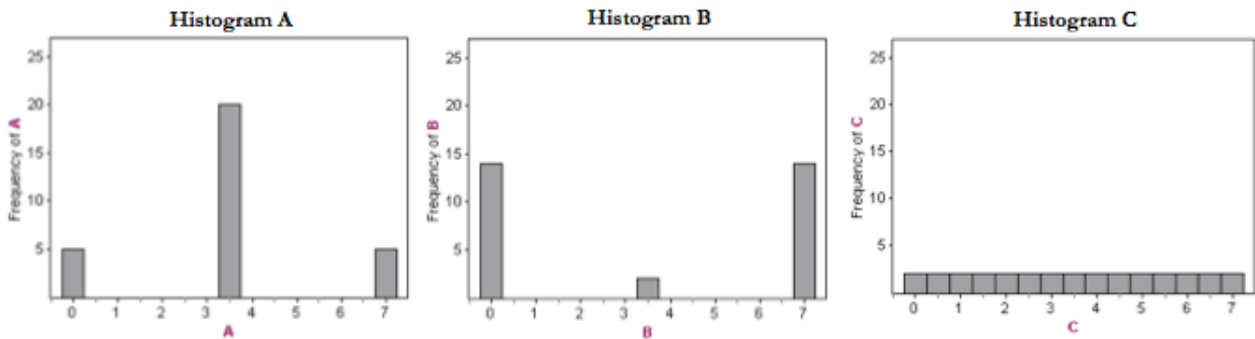


(1) gender, (2) height, (3) sleep, (4) off campus, (5) piercings, (6) time to class

Pay attention to natural boundaries. If skewed, think about whether lower or higher values are more likely.

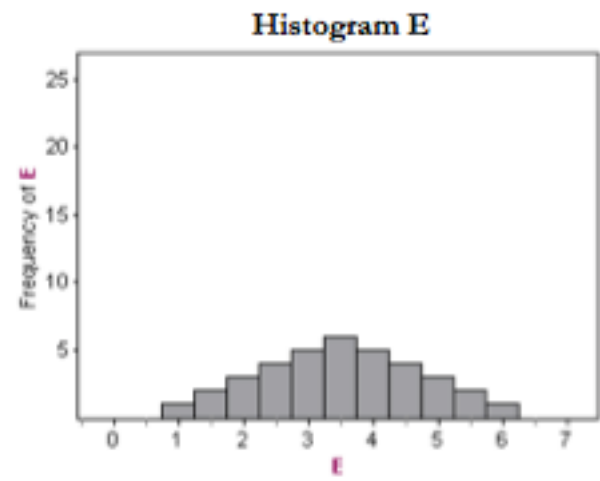
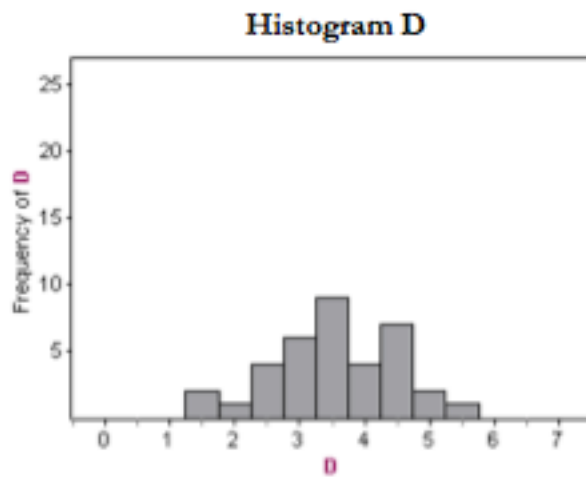
## Variability

3. Order histograms A, B, and C from least to most variable. Explain your reasoning.



A - least, C - medium, B - most

4. Between histograms D and E, which exhibits more variability? Explain your reasoning.



*E more variable, more observations away from mean*