

# DAP: Day 2

## Categorical Variables



*"But where shall I begin?" asked Alice. "Begin at the beginning," the king said gravely, "and go on till you come to the end: then stop."  
-Lewis Carroll, Alice's Adventures in Wonderland*

# Three Rules of Data Analysis

1. **Make a Picture**—A display of your data will reveal things you are not likely to see in a table of numbers and help you to *think* clearly about the patterns and relationships that may be hiding in your data.

2. **Make a Picture**—A well-designed display will *show* the important features and patterns in your data. A picture will also show you the things you did not expect to see: the extraordinary (possibly wrong) data values or unexpected patterns.

3. **Make a Picture**—The best way to *tell* others about your data is with a well-chosen picture.

**Categorical Data**



# Lets Make a Picture

## Categorical Variables and their Graphs

- **Frequency Table (counts)**
- **What is the variable?**
- **Is it quantitative or categorical?**
- **How do we create the proportion and percentage reported in the table?**

**Categorical Data**



- Frequency Table (counts)
- What is the variable?
- Is it quantitative or categorical?
- How do we create the proportion and percentage reported in the table?

## Frequency of Shark Attacks in Various Regions 1996-2006

Region	Frequency	Prop	Percent
Florida	365	0.385	38.5
Hawaii	60	0.063	6.3
California	40	0.042	4.2
Australia	94	0.099	9.9
Brazil	66	0.070	7.0
S.Africa	76	0.080	8.0
Reunion Island	14	0.015	1.5
New Zealand	18	0.019	1.9
Japan	4	0.004	0.4
Hong Kong	6	0.006	0.6
Other	206	0.217	21.7
Total	949	1.00	100

Monday, May 24, 2010

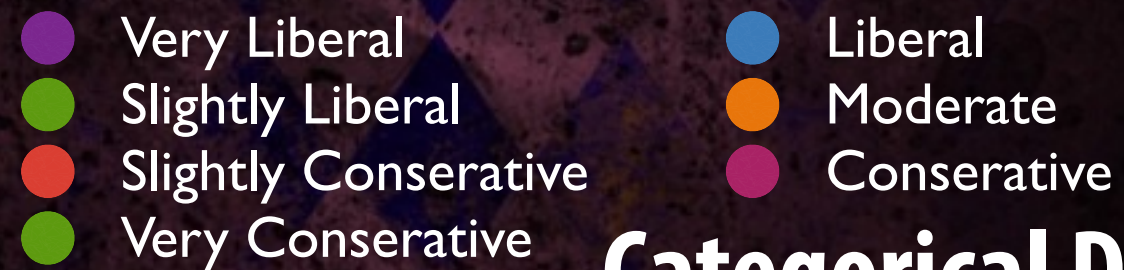
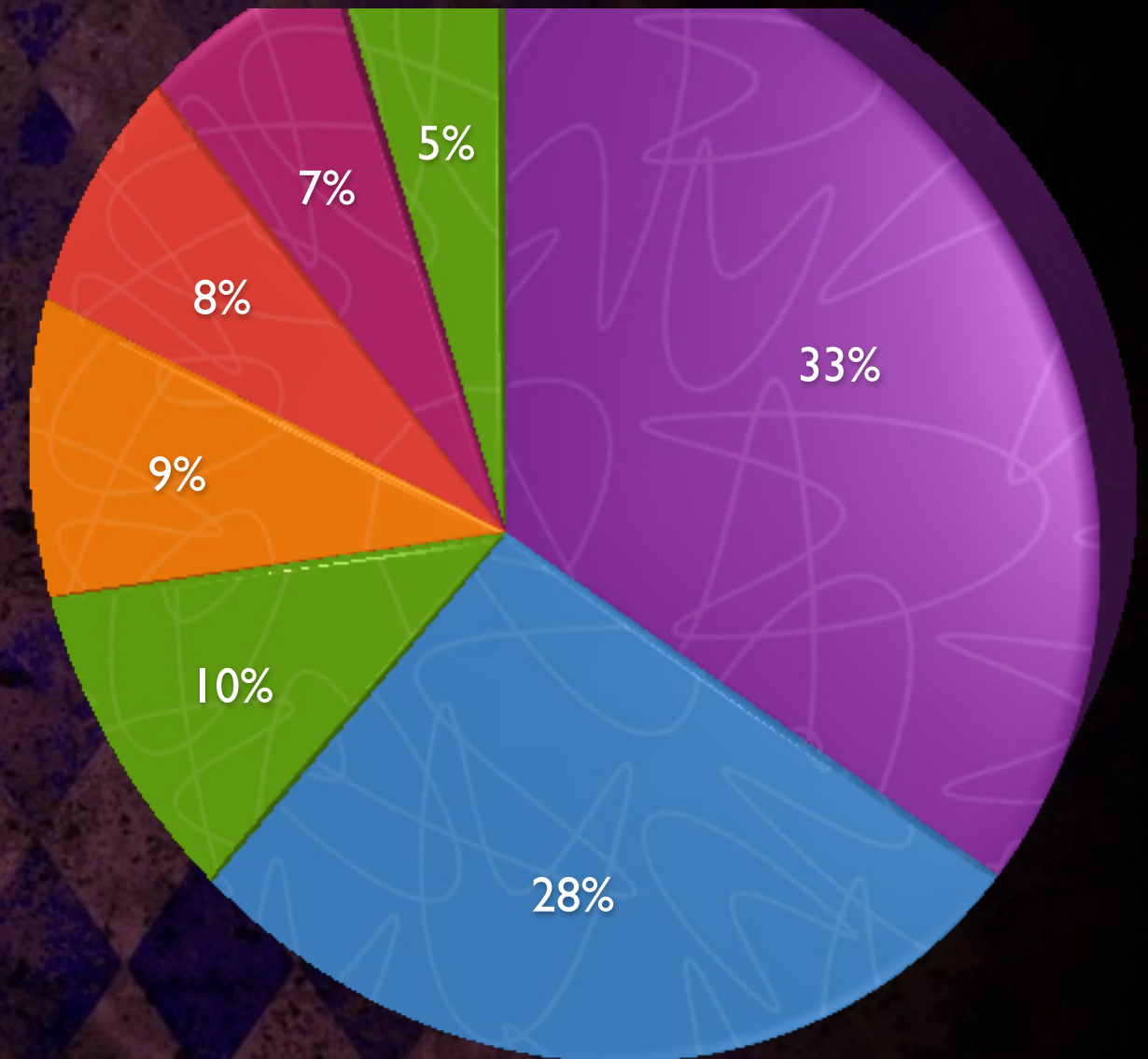
Categorical Variables and their Graphs

- Frequency Table (counts)
- What is the variable?
- Is it quantitative or categorical?
- How do we create the proportion and percentage reported in the table?

# Lets Make a Picture

- Pie Charts
- This type of chart always gives a good visual of the parts to the whole.
- To find the central angle divide the percent by 360.

How's your political view?

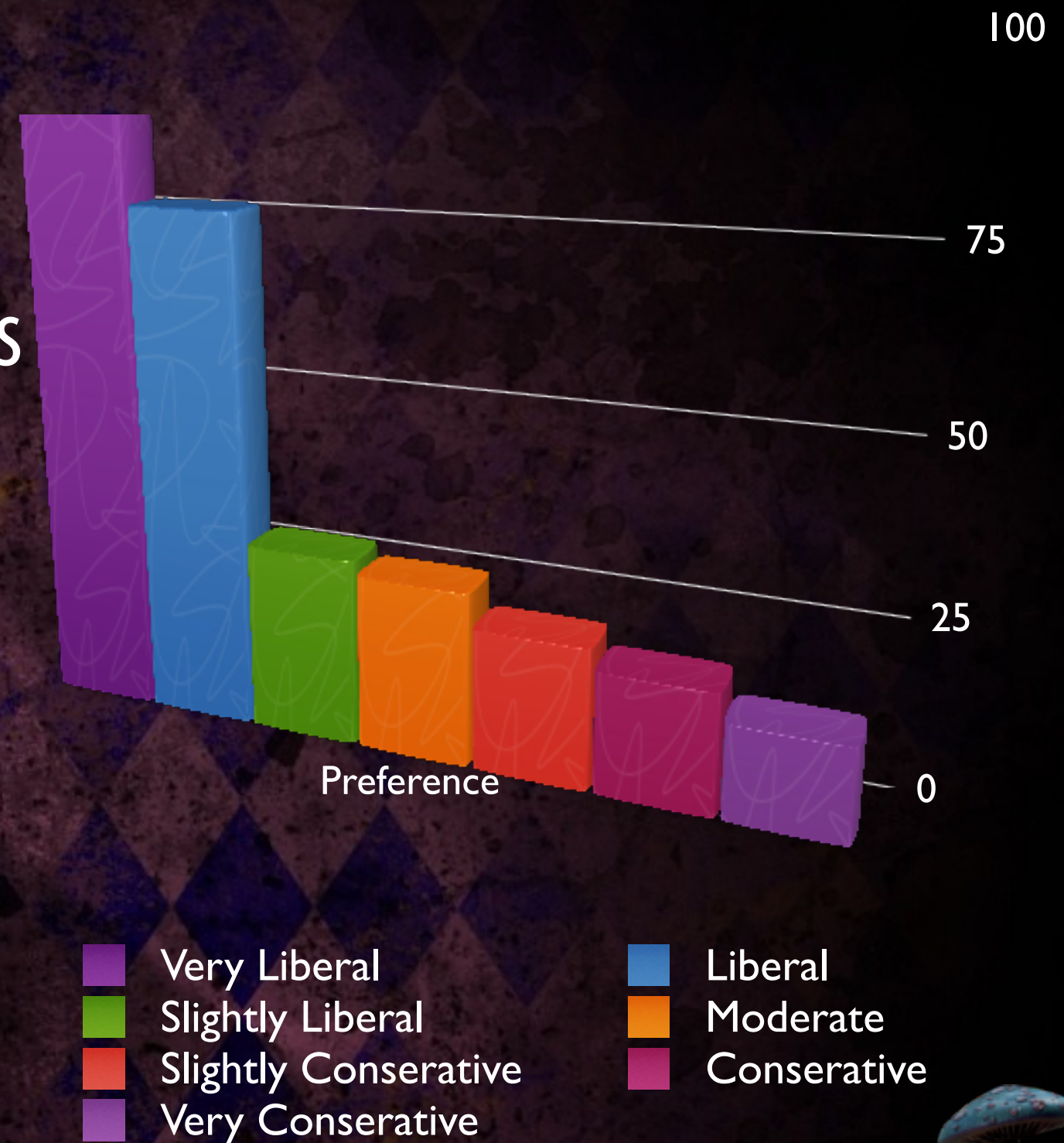


**Categorical Data**



# Lets Make a Picture

- Bar Chart
- Bar charts don't have a necessary order. This type where the largest is first and then put in order to smallest is called a pareto chart.
- The x-axis variable can be put in any order.
- The y-axis variable is the frequency or relative frequency (%) for each category.

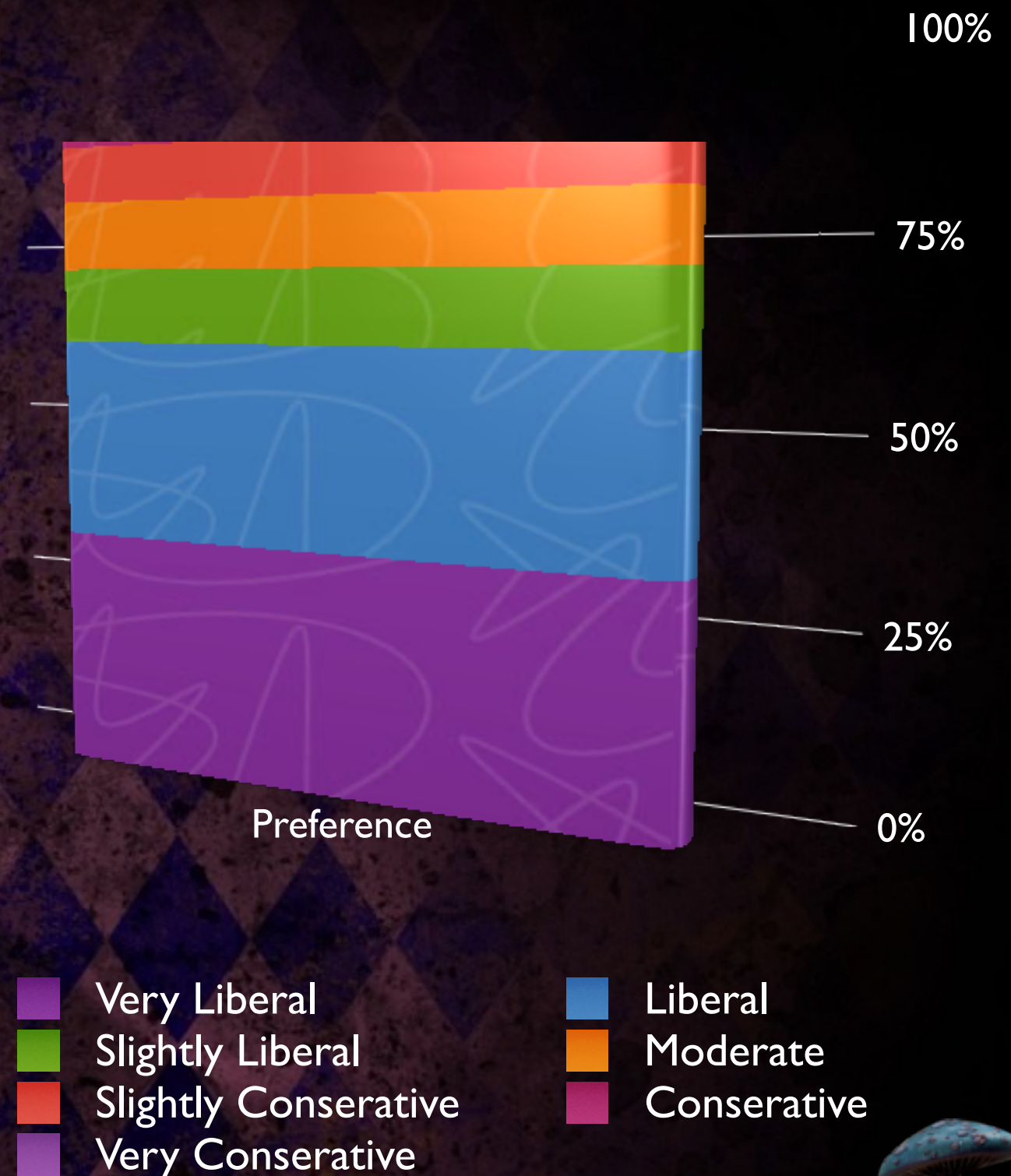


**Categorical Data**



# Lets Make a Picture

- Segmented or Stacked-Bar Chart
- This type of bar chart can bring the benefits of the pie chart into the bar chart.
- The segmented bar chart always has 100% of the data. It is divided up based on percentage or proportions rather than actual counts.



**Categorical Data**



# Lets Make a Picture

- Lets look at the data we collected yesterday.
- Which questions used categorical data?
- Lets use the data in question 14 to create a pie chart and a bar chart or segmented bar chart

**Categorical Data**

