

Python-pptx

Programmatically creating slide decks

Elaine Wilcox-Cook
Data Analyst
September 23, 2019

Python-pptx

- ▶ Package in Python that allows for the programmatic creation of slide decks
- ▶ Creates slides using slide layouts in a slide master
- ▶ This presentation was created entirely using a slide master and python-pptx

Why use Python-pptx?

- ▶ Excellent tool for automated creation of repetitive slides with similar formatting (e.g. 25 slides of bar charts)
- ▶ Create multiple slide decks with no scale-up (e.g. focus group results from different sites)
- ▶ Easily update slide decks that must be created every year (e.g. longitudinal research)
- ▶ Edit formatting automatically when changes need to be made (e.g. change the data labels for all bar charts in presentation)

Creating content

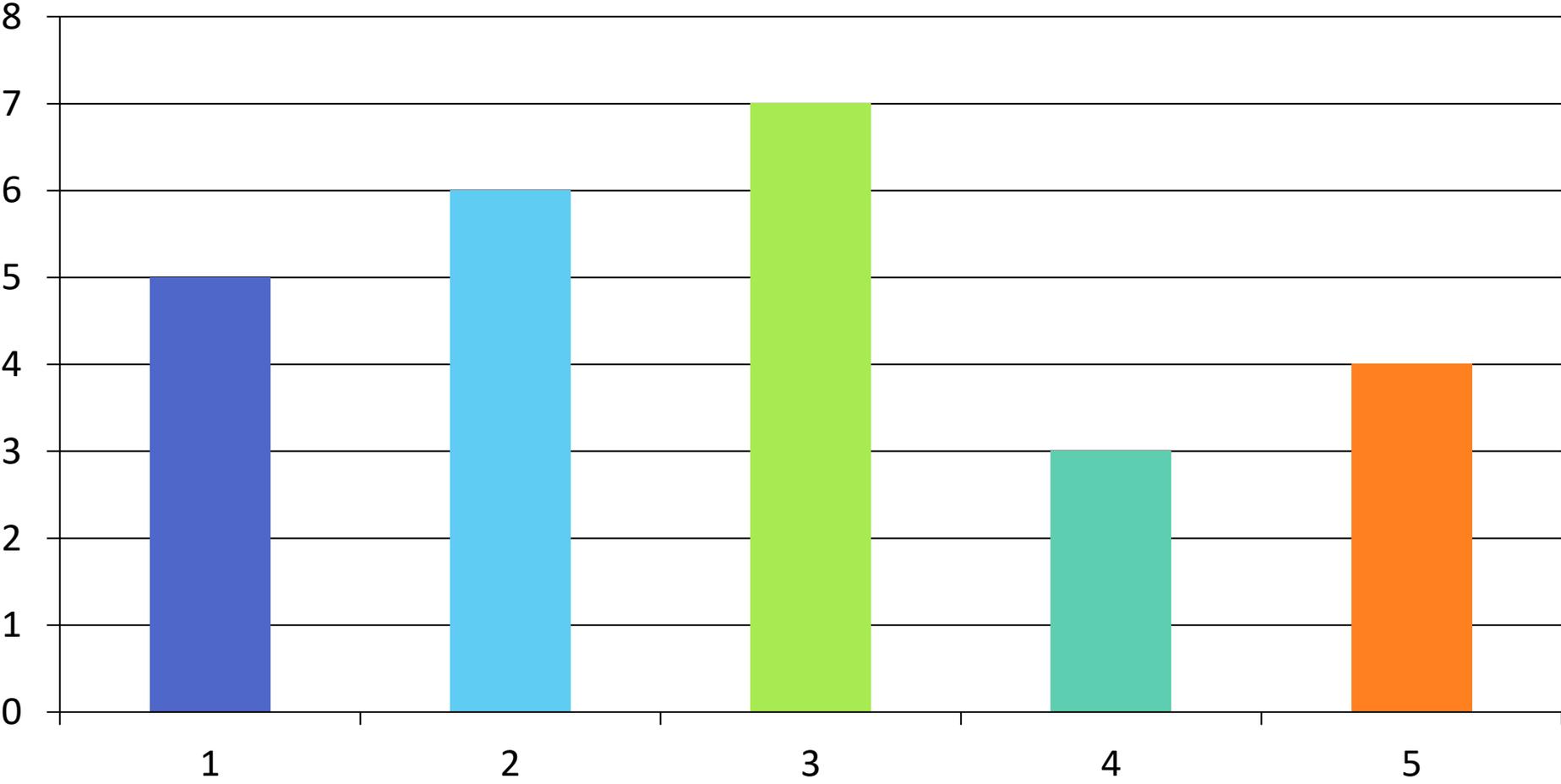
- ▶ Layouts are imported from the slide master of an existing PowerPoint presentation.
 - ▶ If no PowerPoint presentation is specified, the package supplies a default layout
- ▶ Layouts contain placeholders that may be pre-formatted to a certain degree in the slide master. You can insert content directly into a placeholder.
- ▶ Alternately, content can be inserted into a slide without placeholders, but this presentation will not cover that.

Creating content (continued)

- ▶ Placeholder types include:
 - ▶ Textboxes (such as this)
 - ▶ Charts
 - ▶ Images
 - ▶ Tables
 - ▶ And more
- ▶ This presentation will only deal with textboxes, charts, images, and tables

An unpopulated slide with placeholders present

That same slide, populated



Tables

Tables are more complicated to add, since they must be populated cell by cell

Person	Vanilla	Chocolate	Mint	Strawberry
Jamal	●	●	●	
Elise	●	●		●
Anthony	●		●	●
Miriam	●	●		●
Betsy		●	●	●

- = 1st
- = 2nd
- = 3rd

Images

```
#Slide 9
slide=prs.slides.add_slide(image_slide_layout)
title=slide.shapes.title
img_placeholder = slide.shapes[1]

title.text = "Images"
img = img_placeholder.insert_picture('insertimageslide.jpg')
```

Text

- ▶ Textbox content is split into two categories: Paragraphs and Runs
- ▶ A paragraph is all the text contained before a line break, which is created by inserting a new paragraph. A paragraph is made up of runs. A text box always has at least one paragraph, and a paragraph always has at least one run.
- ▶ A paragraph can consist of a single run, or multiple runs. Runs allow **the user to** *vary formatting within a paragraph, a sentence, or even a word.*

Pros/Cons of Python-pptx

Pros

- Good for automation
- Simple, easy to learn
- Flexible

Cons

- Not good with static slides
- Just can't do some things