LIN6209 Coding for Linguists

Weekly Topics

Week 1 - Getting Started

- Python?
- Installing Python
- Starting Python
- Writing Python
- Saving your work
- Variables & functions
 - Creating and using variables
 - Using functions
- Debugging

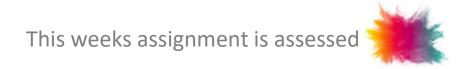
- Objects and object types
 - Integer: int()
 - Floating point: float()
 - Boolean: bool()
 - String (text): str()
- Operators
 - +, -, *, /, //, %
 - <, <=, ==, !=, >=, >
 - and, or, not
 - [], is, in

Week 2 - Functions

- Built-in functions
- Parts of a Function
 - Name
 - Parameters
 - {hidden internal code}
 - Return value
- Defining a function
 - Keyword: def
 - Keyword: pass
 - Keyword: return

- Saving a function
- Re-using a function
 - Keyword: import
 - Keyword: from
- Testing & Debugging
- Good programming style

 The Python standard library and the Python ecosystem



Week 3 — Conditional execution

- Conditional execution
 - Keyword: if
 - Keyword: else
 - Keyword: elif

Boolean logic

Errors & Exceptions

- Iteration
 - Keyword: while
 - Keyword: break
 - Keyword: continue

Testing & Debugging (again)

Week 4 – Collections

- Mutable & Immutable objects
- Collections
 - String
 - List
 - Tuple
 - Range
 - Set
 - Frozenset
 - Dictionary
- Errors & Exceptions

- Iteration
 - Keyword: for
- Comprehensions
- More Collections
 - Queue
 - Deque
 - Stack
 - Tree
 - Graph



Week 5 – Reading and Writing Files

- What a file is
 - A stream of bits
 - Binary, text, CSV ...
 - Records, fields
- Creating, writing, reading, updating, deleting
- Encodings
 - ASCII
 - Unicode

- Keyword: with
 - Always prefer this statement for reading and writing files
- Built-in functions:
 - open()
 - close()
 - know about these but DoNotUse them

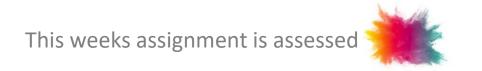
Week 6 – Analysing text (part 1)

- Counting
 - Characters
 - Words
 - Sentences
 - Paragraphs
 - Other structures
- Length
 - Words
 - Sentences
 - Paragraphs

- Adjacency
 - Words
 - Letter sequences

- Other measures
 - Mean, Mode, Median, Variation

Good programming style (again)



Week 7 – Reading Week

- Assignment
 - Analysing some given texts

Week 8 – Regular Expressions

- What a regular expression is
- The RE library
- Meta-characters

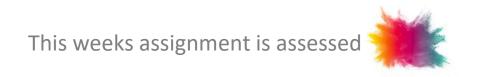
- . Any character except a newline
- [] Character class
- Compliment of class
- \ Backslash
- Zero or more times
- + One or more times
- ? Once or zero times
- {m, n} At least m times and at most n times

- Logical OR
- ^ At the beginning of a line
- \$ At the end of a line
- () Grouping
- \A At the start of a string
- \Z At the end of a string
- **\B** At a word boundary
- \b Inside a word
- (?= ...) Positive lookahead assertion
- (?! ...) Negative lookahead assertion

Week 9 – Analysing text (part 2)

- Using real data
 - gutenberg.org
 - corpusdata.org
 - talkbank.org
 - Beautiful Soup library
- Test data
 - Designing and creating test data
 - www.mockaroo.com
- Data cleansing

- Building software
 - Requirements
 - Analysis
 - Design
 - Build
 - Test
 - Evaluate
- Analysing data
 - Regression analysis
 - Cluster analysis



Week 10 – Data visualisation

- Ready-made libraries
 - matplotlib.org
 - seaborn.pydata.org
 - bokeh.org
 - plotly.com/python
- Publishable graphics
 - Paper
 - Web

- Line graphs
- Charts
- Scatter plots
- Contour maps
- Heatmaps
- Geographic projections
- Animated charts
- Real-time plots

Week 11 – Jupyter Notebooks

- A project notebook
 - A readable document
 - Mixes code, text and graphics
 - Editable
 - Executable
 - Dynamic
 - It makes your work look good

- A record of your work
 - In your browser
 - On your PC
 - Working in teams

Week 12 – Current Trends

- GIT
- Anaconda
- Pandas
- Scikit-learn
- TensorFlow
- ST
- NLTK
- Other languages
- ANTLR

- Big data
- Machine learning
- Artificial intelligence
- Chatbots
- Sentiment analysis
- Opportunities
- Limitations

Each class includes

- PowerPoint slides presented
- Practice exercises
- Weekly assignment
 - A review of the previous weeks assignment
 - This weeks assignment, explained