

Text processing

Dr Peadar Grant

March 13, 2024

Contents

1 Plain text

S.2

2 Text extraction

S.3

1 Plain text

- Can be composed and modified in any text editor
- Most programming languages are written in plain text

Plain text is not:

- Microsoft word documents
- PDF
- Jupyter notebooks

2 Text extraction