

This is a test document

Its *purpose is for integration tests*; To verify that **this text** is correctly extracted.

This is a false title

Ideally it should be picked up by the extractor, even if it is made only using bold and font-size

H1 title

A wee paragraph

H2 title



☐ Some items

☐ More items

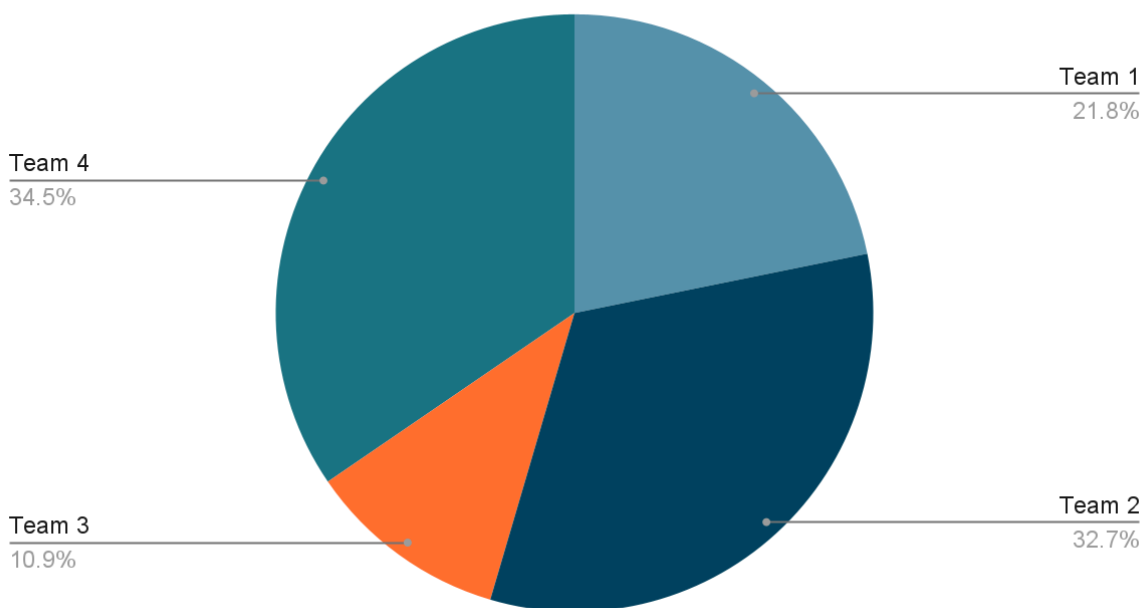
- An unrelated list
- With other items
- Another item
- This is to force
- A page break
- Without using the actual
- Page
- break
- Xml
- Element
- One more
- yay!

This is the second page

OR gate table	true	false
true	true	true
false	true	false

A nice diagram

Points scored



Circle area is πR^2

A forced page break:

My header
ACME INC.

And my table of contents

H1.....	1
H2.....	1
A nice diagram.....	2