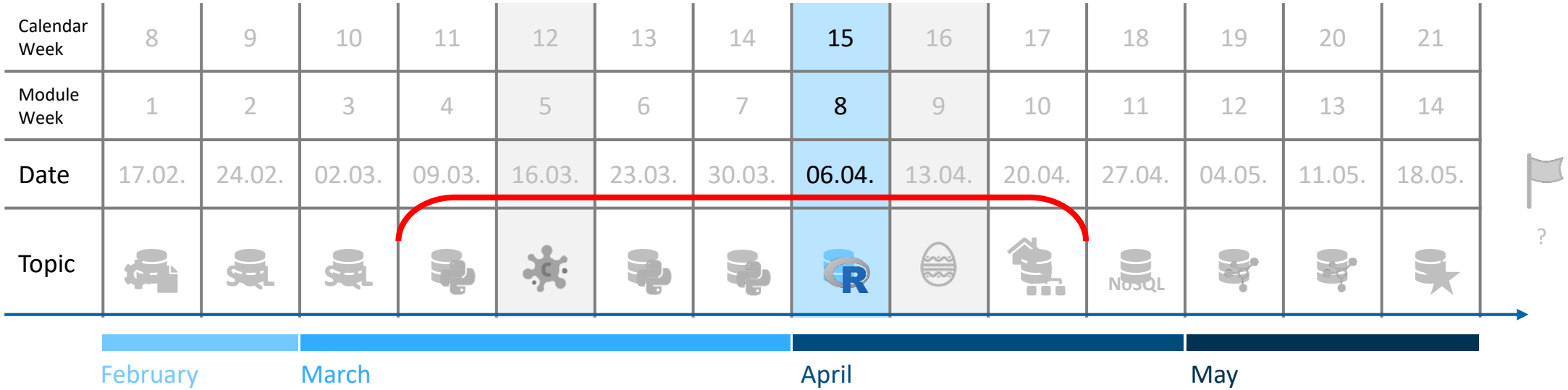
ACLS

R

SS 2020 – Week 8

April 6

Schedule



Introduction & RDB

Structured Query Language

Database & Python

Database & R

Data Warehouse

Not only SQL

Graph Database

Special

Content

• Review		5’
• R and MySQL		40’
• R and SQLite		45’
• Exercises		135’

Content

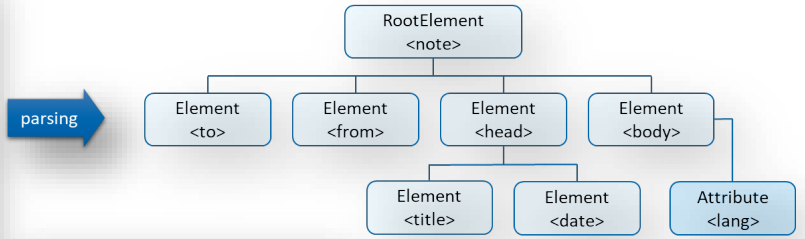
• Review		5’
• R and MySQL		40’
• R and SQLite		45’
• Exercises		135’

Review

Web Scraping


```
<note>
  <to>Tove</to>
  <from>Jani</from>
  <head>
    <title>Reminder</title>
    <date>14/02/2017</date>
  </head>
  <body lang='EN-US'>
    Theater this Sunday!
  </body>
</note>
```

Document



DOM

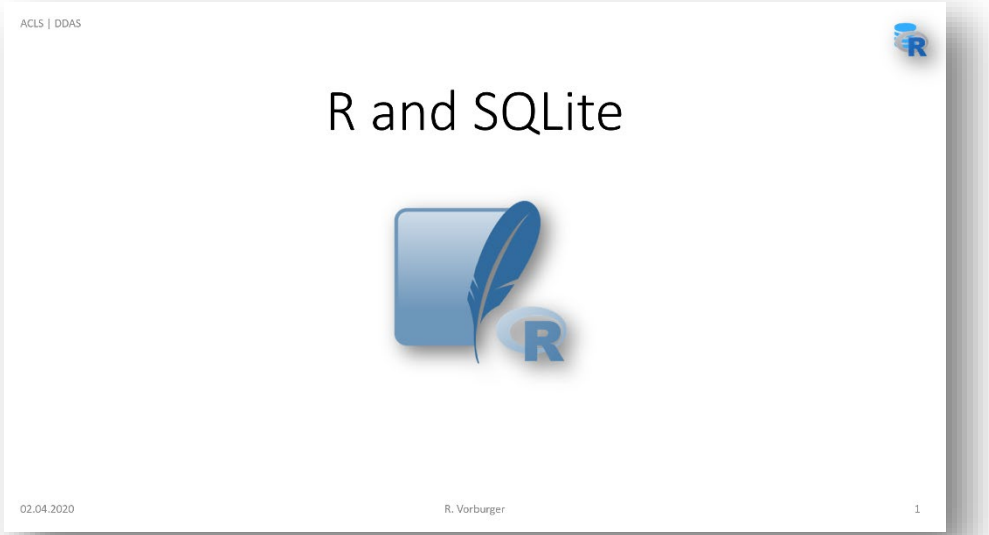

```
# import beautiful soup
import bs4

# get pointer to webpage
wp = urllib.request.urlopen('http://www.somepage.com')

# parse document
soup = bs4.BeautifulSoup(html_doc, 'html.parser')

# get all urls in the document
urls = soup.find_all('a'))
```

Content

• Review		5’
• R and MySQL		40’
• R and SQLite		45’
• Exercises		135’

Content

• Review		5’
• R and MySQL		40’
• R and SQLite		45’
• Exercises		135’



Content

• Review		5’
• R and MySQL		40’
• R and SQLite		45’
• Exercises		135’