Starting the Data Analytics Journey

Data collection

Zeming Yu and Leon Yan
Today’s agenda

• Highs and lows (10 minutes – Leon)
  – Over the next 90 minutes…
  – It’s okay to feel…

• Theoretical case studies (20 minutes – Leon)
  – When’s the best time to fly?
  – How accurate are the punters?

• Real case study (60 minutes – Zem)
  – Collecting Meetup data
  – Analysing Meetup data and offering insights
Act 1

HIGHS AND LOWS
Over the next 90 minutes

Theoretical case studies

Analysing data and offering insights

We are here

Collecting Meetup data
“The only impossible journey is the one you never begin”
It’s okay to feel…

- shock
- denial
- anger
- depression
- acceptance
- integration
Tips

- Get your hands dirty
- Be prepared to work with IT and to debug PC issues
- Choose a pet project
- Build momentum
- Avoid excess multi-tasking
- Tutorials
- Try different packages
- Google for answers
- Collaborate with colleagues
- Write notes and comment code
- Do take breaks
Act 2

THEORETICAL CASE STUDIES
When’s the best time to fly?

17% of industry flights were adversely delayed by 15 minutes or longer
# How accurate are the punters?

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Odds</th>
<th>Implied Probability</th>
<th>Comparison</th>
<th>Perceived Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arsenal wins</td>
<td>1.80</td>
<td>56%</td>
<td>&gt;</td>
<td>45%</td>
</tr>
<tr>
<td>Draw</td>
<td>3.00</td>
<td>33%</td>
<td>&lt;</td>
<td>35%</td>
</tr>
<tr>
<td>Arsenal loses</td>
<td>5.50</td>
<td>18%</td>
<td>&lt;</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>107%</td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>
(...time to get up and stretch)

INTERMISSION
Act 3

REAL CASE STUDY
Case Study
Application Programming Interface

TripView Lite
TripView My Ltd  Travel & Local

Add to Wishlist  Install

TripView displays Sydney and Melbourne public transport timetable data on your phone. It features a summary view showing your next services, as well as a full timetable viewer. All timetable data is stored on your phone, so it can be used offline.

Features:
- Timetable and service interruption information
Flexible, RESTful access to the user's inbox

Gmail API overview

Read and send messages, work with labels, and search for specific threads.
API Reference

This is the class and function reference of scikit-learn. Please refer to the full user guide for further details, as the class and function raw specifications may not be enough to give full guidelines on their uses.

**sklearn.base**: Base classes and utility functions

Base classes for all estimators.

### Base classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>BaseEstimator</code></td>
<td>Base class for all estimators in scikit-learn</td>
</tr>
<tr>
<td><code>Biclustering</code></td>
<td>Biclustering class for all bicluster estimators in scikit-learn</td>
</tr>
<tr>
<td><code>ClassifierMixin</code></td>
<td>Mixin class for all classifiers in scikit-learn</td>
</tr>
<tr>
<td><code>ClusterMixin</code></td>
<td>Mixin class for all cluster estimators in scikit-learn</td>
</tr>
<tr>
<td><code>DensityMixin</code></td>
<td>Mixin class for all density estimators in scikit-learn</td>
</tr>
<tr>
<td><code>RegressorMixin</code></td>
<td>Mixin class for all regression estimators in scikit-learn</td>
</tr>
<tr>
<td><code>TransformerMixin</code></td>
<td>Mixin class for all transformers in scikit-learn</td>
</tr>
</tbody>
</table>

### Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>base.clone(estimator[, safe])</code></td>
<td>Constructs a new estimator with the same parameters.</td>
</tr>
<tr>
<td><code>get_config()</code></td>
<td>Context manager for global scikit-learn configuration</td>
</tr>
<tr>
<td><code>set_config(**kwargs)</code></td>
<td>Set global scikit-learn configuration</td>
</tr>
</tbody>
</table>
API Documentation

TensorFlow has APIs available in several languages both for constructing and executing a TensorFlow graph. The Python API is at present the most complete and the easiest to use, but other language APIs may be easier to integrate into projects and may offer some performance advantages in graph execution.

A word of caution: the APIs in languages other than Python are not yet covered by the API stability promises.

- Python
- C++
- Java
- Go
- Swift (Early Release)

We also provide the C++ API reference for TensorFlow Serving:

- TensorFlow Serving

We encourage the community to develop and maintain support for other languages with the approach recommended by the TensorFlow maintainers. For example, see the bindings for: C#, Haskell, Julia, Ruby, Rust, and Scala.
Application Programming Interface

Extend your community

Meet the API

The Meetup API provides simple RESTful HTTP and streaming interfaces for exploring and interacting Meetup platform from your own apps.

Composing a request

The API is a set of core methods and a common request format. They are combined to form a URL that returns the information you want. Here’s an example of an API call that
Getting an API Key

Hello, Zeming Yu, click the link below to reveal your API key

`-----------------------------`

You’ll need to provide this key with every request you make to the Meetup API. Please try to keep this key to yourself. If you need to reset your key for any reason, you can do that below. Careful, resetting the key will cause any requests made with the old key to fail.

Don’t give away your API key. Your key exposes the Meetup groups you’ve joined, even private groups. It’s personal and belongs to you, like your account password. If you do want a third party service to be able to use your key temporarily and give them the key, you can reset it by clicking the button below. Once reset, your old key is invalidated and no longer works with our API.

[Reset my API key]
https://api.meetup.com/find/groups?&sign=true&photo-host=public&country=Australia&upcoming_events=1&text=actuarial&page=100
https://www.meetup.com/Actuaries-Social-Networking-Group/

What we're about

It's been said that more than 60% of jobs are landed through networking rather than traditional job application methods. Expensive MBA programs are sought after for the value of its alumni network more so than the context of the program. As well as improving job prospects or accessing business opportunities, networking can provide the opportunity to:

- create new relationships and build on existing relationships;
- share knowledge and ideas;
- gain and provide advice and support.

The New Social Networking Committee Goals

The newly formed Social Networking Group’s (SNG) aims to provide Actuaries Institute members with opportunities to network and connect with each other as well as those in non-actuarial but related communities. We plan to hold events that:

- focus on building skills;
- as well as social events that allow members to expand their network.

All our events will be relaxed and accessible and there will be plenty of opportunities for members to participate.

Members (250)

Organized by
Social Networking Actuaries Group and 14 others

Upcoming Meetups

13
Fri, Jul 13, 2018, 5:30 PM
Post HBS & HBS Alumni Results Drinks - SYDNEY

Hosted by Bill Constantinides and 4 others

Attend

28
Fri, Sep 28, 2018, 6:30 PM
Post CAP course result drinks - SYDNEY

Hosted by Kim

Attend

Past Meetups (12)

21
Thu, Jun 21, 2018, 5:30 PM
Post Part III exam results drinks - SYDNEY

Hosted by Highrise

Attend
Next Meetup

**Pre EOFY Drinks - MELBOURNE**

Hosted by Melissa Tan and Jan Le.

Join us for drinks at Whitshart Bar and enjoy the calm before the EOFY storm.

What we're about

It's been said that more than 80% of jobs are landed through networking rather than traditional job application methods. Expensive MBA programmes are sought after for the value of its alumni network more so than the content of the program. As well as improving job prospects or accessing business opportunities, networking can provide the opportunity to:

- create new relationships and build on existing relationships;
- share knowledge and ideas;
- gain and provide advice and support.

The New Social Networking Committee Goals

The newly formed Social Networking Group (SNG) aims to provide Actuaries Institute members with opportunities to network and connect with each other.
JavaScript Object Notation

- Plain text format
- De facto data format for web and mobile apps
- Very flexible - great for unstructured data
- Supported by most programming languages
"0": {
    "name": "Social Networking Actuaries Group",
    "urlname": "Actuaries-Social-Networking-Group",
    "members": 250,
    "next_event": {
        "id": "251221735",
        "name": "Pre EOFY Drinks - MELBOURNE"
    }
},

"1": {
    "name": "Eastern Suburbs Property Investment Network",
    "urlname": "Eastern-Suburbs-Property-Investment-Strategy",
    "members": 569,
    "next_event": {
        "id": "251744354",
        "name": "FREE 1 Hour Property Coaching Session"
    }
},

"2": {
    "name": "North Sydney Property Investment Network",
    ...}
• Key: “name”
• Value: “Social Networking Actuaries Group”
- Key: “urlname”
- Value: “Actuaries-Social-Networking-Group”
- Key: “0”
  - Value: nested JSON object

- Key: “1”
  - Value: nested JSON object

- Key: “2”
  - Value: nested JSON object
Powershell

$url = 'https://api.meetup.com/find/groups?&sign=true&photo-host=public&country=Australia&upcoming_events=1&text=actuarial&page=100&key=INSERTYOURKEY'

Invoke-WebRequest $url |
ConvertFrom-Json |
format-table urlname, name, description |
tee 'd:\output.txt'

PS D:\> Invoke-WebRequest $url |
>> ConvertFrom-Json |
>> format-table urlname, name, description |
>> tee 'd:\output.txt'

<table>
<thead>
<tr>
<th>urlname</th>
<th>name</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actuaries-Social-Networking-Group</td>
<td>Social Networking Actuaries Group</td>
<td>&lt;p&gt;It’s been said that more than 80%...</td>
</tr>
<tr>
<td>Eastern-Suburbs-Property-Investment-...</td>
<td>Eastern Suburbs Property Investment ...</td>
<td>&lt;p&gt;&lt;h&gt;WELCOME&lt;/h&gt;&lt;/p&gt;...</td>
</tr>
<tr>
<td>Little-Bay-Property-Investment-Network</td>
<td>Little Bay Property Investment Network</td>
<td>&lt;p&gt;Live in or around Little Bay? Wan...</td>
</tr>
</tbody>
</table>
Jupyter Notebook Demonstration

This notebook will be published on [Actuaries Digital](http://ActuariesDigital.com) as part of the Analytics Snippet series.
Association analysis (1)

Support: How popular an item set is

Support \{\text{\{\text{apple}\}}\} = \frac{4}{8}

Association analysis (2)

**Confidence.** This says how likely item Y is purchased when item X is purchased, expressed as \( \{X \rightarrow Y\} \).

\[
\text{Confidence} \{\text{Apple} \rightarrow \text{Beer}\} = \frac{\text{Support} \{\text{Apple, Beer}\}}{\text{Support} \{\text{Apple}\}}
\]
Association rules (3)

**Lift**: How likely item Y is purchased when item X is purchased, while controlling for how popular item Y is.

\[
\text{Lift } \{\text{apple} \rightarrow \text{beer} \} = \frac{\text{Support } \{\text{apple, beer} \}}{\text{Support } \{\text{apple} \} \times \text{Support } \{\text{beer} \}}
\]
Zeming Yu
Actuarial Manager
8907 5209
zeming.yu@covermore.com

LinkedIn: www.linkedin.com/in/zeming-yu/

Leon Yan
Actuarial Manager
8907 5219
leon.yan@covermore.com

LinkedIn: www.linkedin.com/in/leon-yan/