🎵Predicting Song Composition with Spotify Data 🎵

Central London Data Science

Wifi:welovecode

Collecting data

Research datasets

- Well curated, great for prototyping or benchmarking
- Already labelled
- Too sterile- May not be representative of real life

Curate your own:

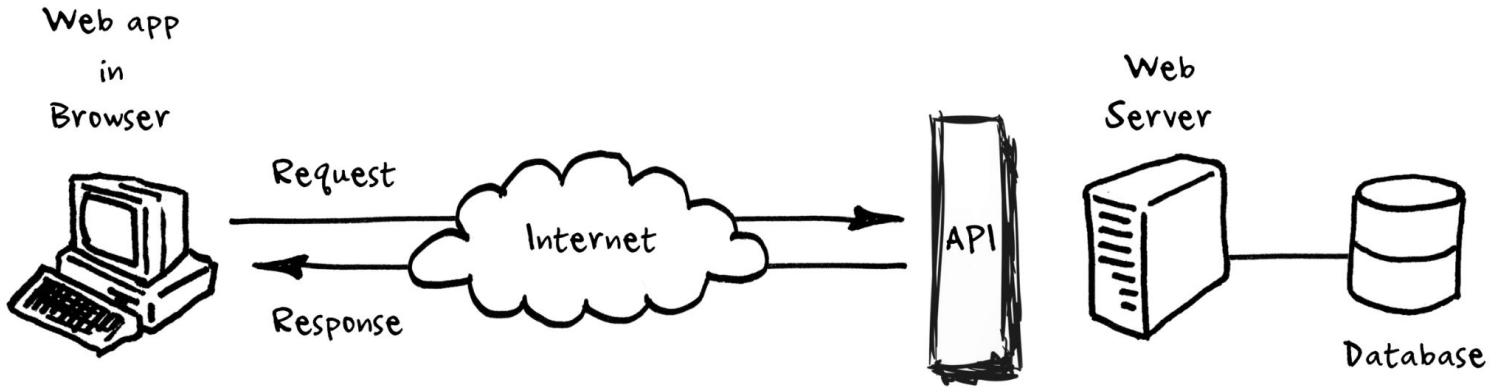
- Arduous
- Expensive

Collecting data

Why not get someone else to do it?

- Companies are collecting as much data as they can
- Significant amount of effort in curating it for you
- It will be real world events

Application Programming Interfaces



Types of music

What is a genre?🧐

- Hiphop, Metal, Acid jazz…
- Is it a group of people or culture that surround a type of music and label it accordingly?
- Or is it music that sounds distinct to any other type of music?

Let’s just stick with the latter tonight….

Timbre

Distinguishes one sound from another

Same amplitude, same frequency same note != same timbre

We use words like:

- Sharp
- Eerie
- Bright

Spotify

- Music streaming service
- Access to millions of songs
- Super rich API

Starting with this...


```
{'meta': {'analyzer_version': '4.0.0', 'platform': 'Linux', 'detailed_status': 'OK', 'status_co
de': 0, 'timestamp': 1444833490, 'analysis_time': 21.35246, 'input_process': 'libvorbisfile L+R
44100->22050'}, 'track': {'num_samples': 11539500, 'duration': 523.33333, 'sample_md5': '', 'of
fset_seconds': 0, 'window_seconds': 0, 'analysis_sample_rate': 22050, 'analysis_channels': 1,
'end_of_fade_in': 0.0, 'start_of_fade_out': 523.33331, 'loudness': -6.82, 'tempo': 125.152, 'te
mpo_confidence': 0.078, 'time_signature': 4, 'time_signature_confidence': 0.877, 'key': 11, 'ke
y_confidence': 0.738, 'mode': 1, 'mode_confidence': 0.664, 'codestring': 'eJxVm1eC4zgSRK-iI8Cb-
19s3wuwZmZ_uouQBIJpIiMN9xzlrF7Wr_xqLaevc9tv1PVr9bZVVuv8NcZvtjXabfXXWym_ede6tc7-G5MPx6xjl37Wb7VR
vOztzHV_7Dx-e5yy5-jnt085v9ZW3Vy132mT-_DPGrv23y3nfJetzt_te_3G8r791l8tc5XfmfXcOSvXdZ7xu23v3Su_rpy
reH3GbrVx3Wf78Qh3tjsO14N_OUc95R6etrOh18tb8HmfJfutfeq8XJ_G2c_kifZh_zHbyf3KKY3fj8v53v2qv5-tc9fFXf
pl5zqH_5bZ776z-IXNsXy0woPz1Xn7QpK7IK7NDXhuJMu9xhmda56r5bqXiiw5d833W0HAXvPXD9Hes3bbv3rrOsp-8QiNz
-_wAQpSvWdy_-v9lH47vSCaUlU1560cwWvvv0_j1-Vsrs9dv3PmRqmbz-viAfl9Z8fh9eb3fH_Mw134BMHwL8bEP9W_esyo
79Iqmu2FXRBYnSiB3_f5j4Iaz9-wvSiw3qWAm0d7Al4qqI2zBgJFzGvnhhOZY11zHA4xuN5oiBueU_q9vxbreze8le8vLYh
jFoRa-HwXH3jxa02fawwhFnDZY3Ktwj4No5C27zqfRWhR7egn7DV70-La-VNgL93PLzb5H4W12znlvWtzroKAuclSYWfVMu
4PNWu3e87ZsXuuJwJGeh1Vr8X14X5cTzWzfx2zO781-HHdS89Es7-uubNH51oFDu411ujz13sbcbDTc3LUE7XpcP34-UY-s
3H_dSf7jSh0rn3LYrvxxLmuHuPHe37iWVqa_vOPA-KaHTMYOthBPMNrcODHk_DclRv3XT8HvusO9t9zne_79XKcfcSDeUu7
6KUfb394srM4ItdqcwxM5XDQX79FJysDb8Nz-MJtsZe1eK4FYuCW-i04VerpXmtwQNkeVcn3OMgSHTBQ9o5DADYFDwZhRlV
_6LNy5Da51gGew865ueYG7qe9IaCB3k_0L5byeRs-x-oc1ZOMpgPjQKhZsBwNw-W6FzxzoLleY29ct55rxXw5BnDRdF31DV
4jgbNB4eF5bwFYwFnuryR_uMrEJmZguZVn332BcGN-iBf7FsVvzvNn39hg-S_iDfyi_QfxxsaA_qNw7OEo8MnlEGJHAJ4vI
```

Turning it into something that looks like this...