

# IKDD Project - 第八組

## Final Project Report

資訊四 鄭宇軒

資訊四 林宏誠

# 大綱

- ▶ 主題簡介 - YouTube Funny Video Chooser
- ▶ 工具 – Youtube Data API
- ▶ 數據分析
- ▶ 演算法與實踐
- ▶ 案後檢討

# 簡介 - YouTube Funny Video Chooser

- ▶ Youtube – 最知名.大量的影片分享網站
- ▶ UCI - YouTube Comedy Slam Preference Data Data Set , 包含約90萬筆的影片比較資料
- ▶ 有趣影片的**共通性**?
- ▶ Youtube有趣影片自動分辨程式 - 判斷兩影片何者較為有趣
- ▶ Github : <https://github.com/ls2jsc/IKDD-Project>

# 工具 - Youtube Data API

- ▶ 功用：取得與影片相關的統計資料
- ▶ 最新版本：Version 3.0，提供 by Google Developers
- ▶ 此專案使用Version 2.0
- ▶ 範例網址：
  - ▶ <http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2&alt=json> (JSON)
  - ▶ 上述網址回傳json檔, 裡面可取得如觀看數、喜歡人數、影片標籤等資訊
  - ▶ <http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments?v=2&max-results=50&orderby=published>
  - ▶ 上述網址回傳該影片最近的50個評論的內容

# Json格式(影片資料)

```
{
  "version": "1.0",
  "encoding": "UTF-8",
  "entry": {
    "xmlns": "http://www.w3.org/2005/Atom",
    "xmlns$gd": "http://schemas.google.com/g/2005",
    "xmlns$yt": "http://gdata.youtube.com/schemas/2007",
    "xmlns$media": "http://search.yahoo.com/mrss/",
    "gd$etag": "W/\"DEUCQX47eCp7I2A9XRRS5GEU.\"\"",
    "id": {"$t": "tag:youtube.com,2008:video:iaosnMCWli8"},
    "published": {"$t": "2014-12-08T13:14:30.000Z"},
    "updated": {"$t": "2015-01-22T00:24:20.000Z"},
    "category": [{"scheme": "http://schemas.google.com/g/2005#kind", "term": "http://gdata.youtube.com/schemas/2007#video"}, {"scheme": "http://gdata.youtube.com/schemas/2007/categories.cat", "term": "Entertainment", "label": "娛樂"}],
    "title": {"$t": "爆笑超短片大合集22!"},
    "content": {"type": "application/x-shockwave-flash", "src": "http://www.youtube.com/v/iaosnMCWli8?version=3&f=videos&app=youtube_gdata", "link": [{"rel": "alternate", "type": "text/html", "href": "http://www.youtube.com/watch?v=iaosnMCWli8&feature=youtube_gdata"}, {"rel": "http://gdata.youtube.com/schemas/2007#video.related", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/iaosnMCWli8/related?v=2"}, {"rel": "http://gdata.youtube.com/schemas/2007#mobile", "type": "text/html", "href": "http://m.youtube.com/details?v=iaosnMCWli8"}, {"rel": "http://gdata.youtube.com/schemas/2007#uploader", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/users/IF_gt4BfsWYm_2G0cKXyEQ?v=2"}, {"rel": "self", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/iaosnMCWli8?v=2"}], "author": [{"name": {"$t": "啾啾鞋"}, "uri": {"$t": "http://gdata.youtube.com/feeds/api/users/chuchushoeTW"}, "yt$userId": {"$t": "IF_gt4BfsWYm_2G0cKXyEQ"}}, {"action": "commentVote", "permission": "allowed"}, {"action": "videoRespond", "permission": "moderated"}, {"action": "rate", "permission": "allowed"}, {"action": "embed", "permission": "allowed"}, {"action": "list", "permission": "allowed"}, {"action": "autoPlay", "permission": "allowed"}, {"action": "syndicate", "permission": "allowed"}], "gd$comments": {"gd$feedLink": {"rel": "http://gdata.youtube.com/schemas/2007#comments", "href": "http://gdata.youtube.com/feeds/api/videos/iaosnMCWli8/comments?v=2", "countHint": 40}}, "media$group": {"media$category": [{"$t": "Entertainment", "label": "娛樂", "scheme": "http://gdata.youtube.com/schemas/2007/categories.cat"}], "media$content": [{"url": "http://www.youtube.com/v/iaosnMCWli8?version=3&f=videos&app=youtube_gdata", "type": "application/x-shockwave-flash", "medium": "video", "isDefault": "true", "expression": "full", "duration": 76, "yt$format": 5}, {"url": "rtsp://r2---sn-jc47eu7r.c.youtube.com/CiILENy73wIaGQkvlpbAnCyqiRMYSARFEgUGz2aWRlb3MM/0/0/0/video.3gp", "type": "video/3gpp", "medium": "video", "expression": "full", "duration": 76, "yt$format": 1}, {"url": "rtsp://r2---sn-jc47eu7r.c.youtube.com/CiILENy73wIaGQkvlpbAnCyqiRMYSARFEgUGz2aWRlb3MM/0/0/0/video.3gp", "type": "video/3gpp", "medium": "video", "expression": "full", "duration": 76, "yt$format": 6}], "media$credit": [{"$t": "chuchushoeTW", "role": "uploader", "scheme": "urn:youtube", "yt$display": "啾啾鞋", "yt$type": "partner"}], "media$description": {"$t": "翻譯粉絲團: https://www.facebook.com/ChuchushoeTrans\nTwitter: https://twitter.com/chuchushoeTW\nnGoogle+: https://plus.google.com/+chuchushoeTW/posts\nnInstagram: http://instagram.com/chuchushoeTW\nn全系列完整播放清單: http://goo.gl/4djV21", "type": "plain"}, "media$keywords": {}, "media$license": {"$t": "youtube", "type": "text/html", "href": "http://www.youtube.com/t/terms"}, "media$player": [{"url": "http://www.youtube.com/watch?v=iaosnMCWli8&feature=youtube_gdata_player"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/default.jpg", "height": 90, "width": 120, "time": "00:00:38", "yt$name": "default"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/mqdefault.jpg", "height": 180, "width": 320, "yt$name": "mqdefault"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/hqdefault.jpg", "height": 360, "width": 480, "yt$name": "hqdefault"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/sddefault.jpg", "height": 480, "width": 640, "yt$name": "sddefault"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/1.jpg", "height": 90, "width": 120, "time": "00:00:19", "yt$name": "start"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/2.jpg", "height": 90, "width": 120, "time": "00:00:38", "yt$name": "middle"}, {"url": "http://i.ytimg.com/vi/iaosnMCWli8/3.jpg", "height": 90, "width": 120, "time": "00:00:57", "yt$name": "end"}], "media$title": {"$t": "爆笑超短片大合集22!", "type": "plain"}, "yt$duration": {"seconds": 76}, "yt$uploaded": {"$t": "2014-12-08T13:14:30.000Z"}, "yt$uploaderId": {"$t": "UCIF_gt4BfsWYm_2G0cKXyEQ"}, "yt$videoid": {"$t": "iaosnMCWli8"}, "gd$rating": {"average": 4.9726496, "max": 5, "min": 1, "numRaters": 585, "rel": "http://schemas.google.com/g/2005#overall"}, "yt$statistics": {"favoriteCount": 0, "viewCount": 132893}, "yt$rating": {"numDislikes": 4, "numLikes": 581}}}
```



# 評論資料(以tag作為分辨要件, 下方紅框內為一條評論的內容)

<?xml version='1.0' encoding='UTF-8'?><feed xmlns='http://www.w3.org/2005/Atom' xmlns:gd='http://schemas.google.com/g/2005' xmlns:openSearch='http://a9.com/-/spec/opensearch/1.1/' xmlns:yt='http://gdata.youtube.com/schemas/2007' gd:etag='W/8quot;CUHfQncycp7I2A9XRRREks.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comments</id><updated>2015-01-26T09:10:33.990Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><logo>http://www.gstatic.com/youtube/img/logo.png</logo><link rel='http://schemas.google.com/g/2005#feed' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments?v=2' /><link rel='http://schemas.google.com/g/2005#batch' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/batch?v=2' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments?max-results=50&orderby=published&v=2' /><link rel='service' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments?alt=atom-service&v=2' /><author><name>YouTube</name><uri>http://www.youtube.com/</uri></author><generator version='2.1' uri='http://gdata.youtube.com'>YouTube data API</generator><openSearch:totalResults>49</openSearch:totalResults><openSearch:itemsPerPage>50</openSearch:itemsPerPage><entry gd:etag='W/8quot;DUEF5H47eCp7I2A9XRRREk.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z123gharsyeziipko04c155onxjkzny5f4c</id><published>2015-01-26T07:33:39.000Z</published><updated>2015-01-26T07:33:39.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>siri那個甚麼醜臉呀XDDDDDD</title><content>siri那個甚麼醜臉呀XDDDDDD</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z123gharsyeziipko04c155onxjkzny5f4c?v=2' /><author><name>蔣 GZB</name><uri>http://gdata.youtube.com/feeds/api/users/BDxFSMvW7AH9Qc6V1mVF\_w</uri></author><channelId>UC8DXFSMvW7AH9Qc6V1mVF\_w</channelId><yt:googlePlusUserId>113245949347053300715</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;DKclRn47eCp7I2A9XRRREU0.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z13nupabx13pj3cp104cgjcatoycwnhwb4k</id><published>2015-01-24T12:54:57.000Z</published><updated>2015-01-24T12:54:57.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>阿嬌好笑了!!!</title><content>阿嬌好笑了!!!</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z13nupabx13pj3cp104cgjcatoycwnhwb4k?v=2' /><author><name>周俊豪</name><uri>http://gdata.youtube.com/feeds/api/users/f6TnGirb8qh6Rtw0JzFWw</uri></author><channelId>f6TnGirb8qh6Rtw0JzFWw</channelId><yt:googlePlusUserId>11679036875644663527</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;CKUDR347eCp7I2A9XRRREU0.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z122dlhakuasupatj2i1l1ohtiwd12d</id><published>2015-01-24T11:51:16.000Z</published><updated>2015-01-24T11:51:16.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>阿嬌那裏哀哀啼啼~~~~</title><content>阿嬌那裏哀哀啼啼~~~~</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z122dlhakuasupatj2i1l1ohtiwd12d?v=2' /><author><name>洪承慈</name><uri>http://gdata.youtube.com/feeds/api/users/VIRN9S5Y6KM2kBVad-1A2zg</uri></author><channelId>UCVIRN9S5Y6KM2kBVad-1A2zg</channelId><yt:googlePlusUserId>114312933503529120392</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;C0YCR347eCp7I2A9XRRSGU4.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z123hridlms1i1mf231ereqixevf2ou04</id><published>2015-01-22T15:39:26.000Z</published><updated>2015-01-22T15:39:26.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>郭最後一個我快笑到斷氣XDD</title><content>郭最後一個我快笑到斷氣XDD</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z123hridlms1i1mf231ereqixevf2ou04?v=2' /><author><name>冷血的猴子</name><uri>http://gdata.youtube.com/feeds/api/users/SR\_RvEEIq3rVkuuKXE\_7DA</uri></author><channelId>SR\_RvEEIq3rVkuuKXE\_7DA</channelId><yt:googlePlusUserId>117510106479231792559</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;DEMBQ347eCp7I2A9XRRSGU4.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z12wcmzatfijnk1322dclcztoeuhd3t</id><published>2015-01-22T14:20:52.000Z</published><updated>2015-01-22T14:20:52.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>原來vines 系列有人在翻譯呀!!! ...</title><content>原來vines 系列有人在翻譯呀!!!不錯呀....標題+個vines 可以讓大家瞭解更多唷~~</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z12wcmzatfijnk1322dclcztoeuhd3t?v=2' /><author><name>田奇盛</name><uri>http://gdata.youtube.com/feeds/api/users/xx654321g</uri></author><channelId>UCVIRN9S5Y6KM2kBVad-1A2zg</channelId><yt:googlePlusUserId>111936614023461494787</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;DKQD5X47eCp7I2A9XRRSGU4.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z13psnk011jellaj204ccvngft35w33pyi0</id><published>2015-01-22T13:46:18.000Z</published><updated>2015-01-22T13:46:18.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>天下同儕都好像大大...</title><content>天下同儕都好像大大...</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z13psnk011jellaj204ccvngft35w33pyi0?v=2' /><author><name>Craft of Duty 22</name><uri>http://gdata.youtube.com/feeds/api/users/INPxNj96vurjPEghLvcSuw</uri></author><channelId>INPxNj96vurjPEghLvcSuw</channelId><yt:googlePlusUserId>103822372301562489201</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;CEQGRH47eCp7I2A9XRRSGU4.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z12nev5bdkulpsnuq230wbxcxsndj5ihb04</id><published>2015-01-22T13:12:05.000Z</published><updated>2015-01-22T13:12:05.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>Which one?</title><content>Which one?</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z12nev5bdkulpsnuq230wbxcxsndj5ihb04?v=2' /><author><name>刺客</name><uri>http://gdata.youtube.com/feeds/api/users/111ewqec</uri></author><channelId>UC\_PtM394klhkKEe06\_gaSDQ</channelId><yt:googlePlusUserId>109389374529235593525</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;A04ESH47eCp7I2A9XRRSGU8.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z12bx10z3ua1grw3m23txfsjgkfseplzu04</id><published>2015-01-22T12:31:49.000Z</published><updated>2015-01-22T12:31:49.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>最後一個也太蝦哈哈哈哈哈</title><content>最後一個也太蝦哈哈哈哈哈</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z12bx10z3ua1grw3m23txfsjgkfseplzu04?v=2' /><author><name>胡荅源</name><uri>http://gdata.youtube.com/feeds/api/users/so32012005</uri></author><channelId>UCJhCeank2ZqLLpNtL4bJAk</channelId><yt:googlePlusUserId>111827888675125295781</yt:googlePlusUserId><yt:replyCount>0</yt:replyCount><yt:videoId>nKcAFrZ-5dc</yt:videoId></entry><entry gd:etag='W/8quot;A04ESH47eCp7I2A9XRRSGU8.&quot;';><id>tag:youtube.com,2008:video:nKcAFrZ-5dc:comment:z12cefc45yjjtrmer22ic3fbf1exdvfk</id><published>2015-01-21T13:06:28.000Z</published><updated>2015-01-21T13:06:28.000Z</updated><category scheme='http://schemas.google.com/g/2005#kind' term='http://gdata.youtube.com/schemas/2007#comment' /><title>求0:33的鋼琴伴奏曲~~~~~!!!!</title><content>求0:33的鋼琴伴奏曲~~~~~!!!!</content><link rel='related' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc?v=2' /><link rel='alternate' type='text/html' href='http://www.youtube.com/watch?v=nKcAFrZ-5dc' /><link rel='self' type='application/atom+xml' href='http://gdata.youtube.com/feeds/api/videos/nKcAFrZ-5dc/comments/z12cefc45yjjtrmer22ic3fbf1exdvfk?v=2' /><author><name>周欣萍</name><uri>http://gdata.youtube.com/feeds/api/users/C3skWk\_LjlfErgtIcdBQ</uri>

# 數據分析

- ▶ 用Python語言去分析 & 兩台電腦去run同一份code
- ▶ 因為時間關係，只parse前面50萬筆的資料。
- ▶ 使用平均數. Normal distribution對評比中有趣/不有趣這二資料群做統計，並以Jaccard Similarity來計算「有趣影片」間相似的要素
- ▶ 此專案中我們用以下幾個可能影響結果的因素作分析：
  - ▶ (1) 觀看數      (2) 喜歡/不喜歡數      (3) (喜歡/不喜歡數)/觀看數
  - (4) 影片種類      (5) 影片時間長短      (6) (影片總觀看數/釋出時間)
  - (7) 評論數      (8) 評論內容



# 遇到的問題

- ▶ 裡面有將近一半的影片代碼為私人或是已經被刪除的影片，導致這些影片的資料無法被取得，最後得到的有效資料合計為248,967筆
- ▶ 來源資料多為2010當年釋出，且皆被歸類為「喜劇」此類別的影片，致使(4)影片種類與(6)影片總觀看數/釋出時間此二因素在training data中無法明顯看出其特徵
- ▶ 試圖從評論挖掘資料時，發覺裡面的語言多而廣泛，此外內容也多為口語化或不同文化所無法理解的流行語

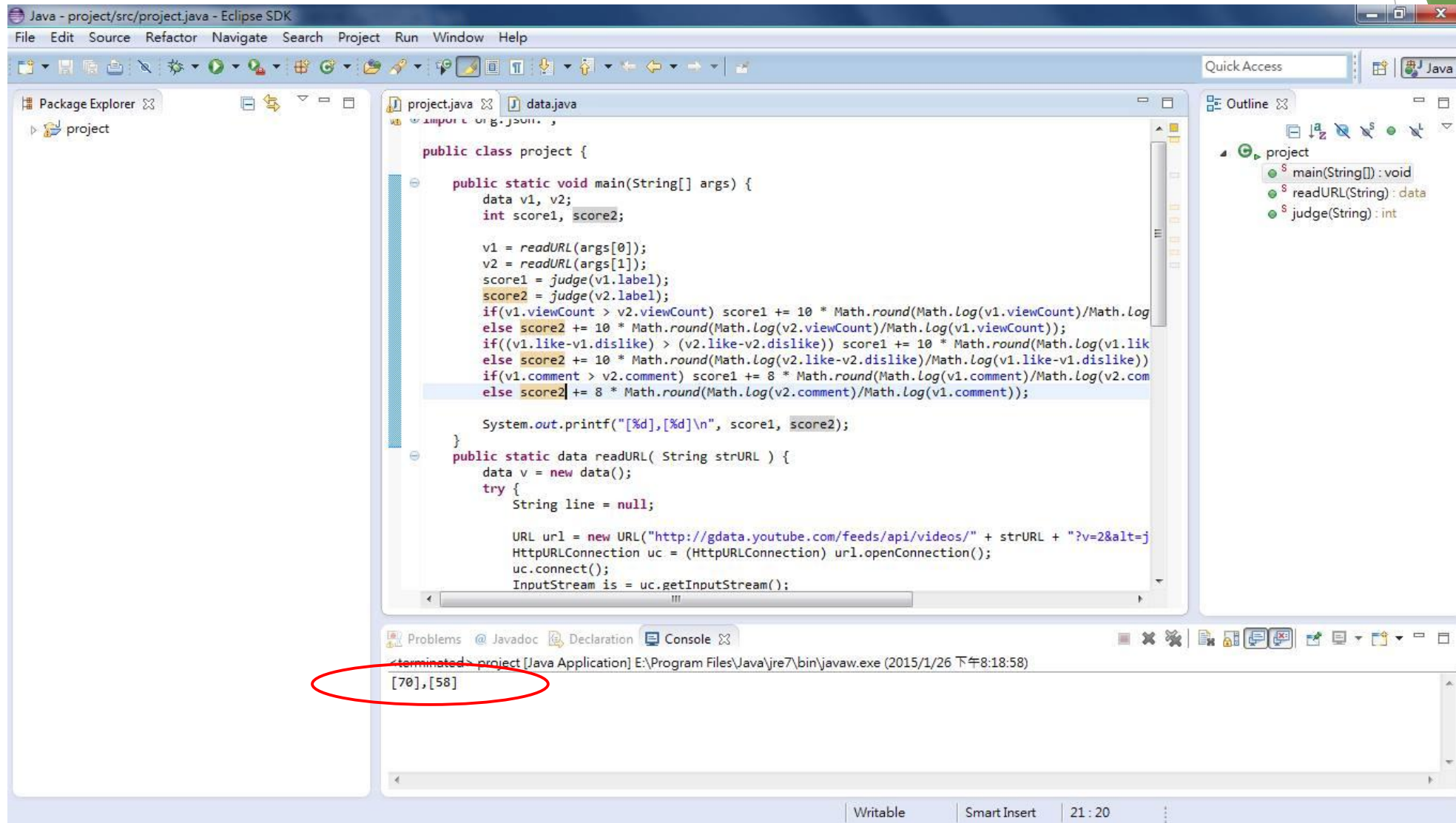


# 演算法與實踐

- ▶ 分析過後，相似度和影響度皆超過50%的因素依其重要度排序有以下四點：
  - ▶ 影片的分類標籤 (「娛樂」, 「喜劇」類別尤其明顯)
  - ▶ 觀看人數(權重:10)
  - ▶ 喜歡/不喜歡(權重:10)
  - ▶ 評論數(權重:8)
- ▶ 概念：
  - ▶ 取得影片在此四重要因素的數據。
  - ▶ 除了“影片標籤”外，以10為底數，對兩影片的比值取其log值，再乘以其權重，即為該影片在此項目所得到的分數

- ▶ 影片標籤(在json中的['entry']['category'][1]['term']) :
  - ▶ 類別為Entertainment或Comedy : +50分
  - ▶ 類別為Film、Animals或Sports : +20分
  - ▶ 其餘 : 0分
- ▶ 觀看人數(在json中的['entry']['yt\$statistics']['viewCount']) :
  - ▶ 數量較多者 :  $+ 10 * \log(\text{view1}/\text{view2})$
- ▶ 喜好與否(在json中的['entry']['yt\$rating']['numLikes']和['numDislikes']) :
  - ▶  $\text{Favor} = \text{Likes} - \text{DisLikes}$
  - ▶ 數量多者 :  $+ 10 * \log(\text{Favor1}/\text{Favor2})$
- ▶ 評論數(在json中的['entry']['gd\$comments']['gd\$feedLink']['countHint']) :
  - ▶ 數量多者 :  $+ 8 * \log(\text{Comments1}/\text{Comments2})$
- ▶ 將以上各項評分加總起來，即為影片最後的得分

# 範例畫面



# 案後檢討

- ▶ 討論過後，覺得在此計畫中有以下幾個缺失：

- ▶ **題目過於單調：**

- 因為計畫的方向只以判斷出兩個影片中哪個比較有趣，二元的單純結果。結果即使寫出程式對此做出判斷，也難以在output中得到更多的資訊。若以“預測影片在幾個月後的閱覽數成長”等具有更多豐富意義的主題為目標的話，可能會有更多有意義的發現。

- ▶ **資料過時與不全：**

- 由於所找到的為2010年的調查資料，致使在數據分析中增添意料外的麻煩。如：無效的影片代碼、至今非一定準確的喜好結果。另外多元語系的評論，也讓我們預期藉由Text Mining收集關鍵詞的目標無法達成。

- ▶ **未引用其他模型為評分依據：**

- 由於時間因素，未能找到相關的模型來套用至此次的研究主題上，故演算方面便是以比較級距的方法來評斷。結果出來的結果不盡預期，可見要自立導出一個合理的公式並非一朝一夕可成之事。



## ► 未來可能發展方向

### ► 影片引用調查：

這次可惜之處在於沒時間運用Python-Scrapy的Web Spider技術，將“影片被引用次數”此項目作為分析參數之一。在社交網路盛行的時代，此一要素極有可能是影響影片熱門與否的要素之一。

### ► 分析特定族群喜好：

雖然在YouTube API中尚未提供觀眾或評論者的年齡層資料，日後若有開放這部分的資料，將大大助益於分析各年齡層影片偏好傾向的研究。

### ► 預估影片閱覽成長：

如果能根據一影片貼出來短時間內的閱覽資料，預測一個影片未來的熱門度/喜好成長的話，在商業運用上將能更有效率地對新推出的行銷影片作評估，進而減少不必要的風險與損失。