**華語流行音樂趨勢分析: 2005-2015**

**廖偉帆[[1]](#footnote-1)、李瑞元**

實踐大學 資訊科技與管理學系碩士班，台北，10462

**摘要** 隨著大數據的時代來臨，數位音樂串流服務，已經改變流行音樂產業；不論是在收聽的客群、喜好、甚至是對熱門金曲預測方式等，都與過去明顯不同。本研究探索近十年間來，華語流行音樂的趨勢發展；藉由網路爬蟲，蒐集KKBOX數位音樂串流服務，自2005年至2015年間，每日華語流行音樂前一百名排行榜；共挖掘30萬幾萬筆資料。數據資料分析發現，前十名的熱門歌曲中，以戲劇主題曲為主流；另外，使用數位音樂串流的聽眾，對於新發行的歌曲有較高的喜好程度。本研究期望能協助數位串流音樂服務，能更了解其聽眾喜好，並提供傳統音樂產業營運之參考。

**關鍵字** 華語流行音樂、趨勢分析、大數據、數位串流音樂、KKBOX

**Chinese pop music Trend Analysis: 2005-2015**

**Wei-Fen Liao1, Maria R. Lee**

Graduate Institute of Information Technology and Management, Shih Chien University, Taipei, 10462

**Abstract** With the advent of the era of big data, music streaming service has changed the popular music industry. There are many significantly changes different from the past, such as target listening groups, preferences, and popular songs predictive mode etc. This study explores the music industry's development trend in Taiwan in this decade. The study crawls daily top one hundred Chinese pop music in KKBOX music steaming from year 2005 to year 2015. More than 300,000 songs are collected. Data analysis shows that the top ten popular songs are drama theme songs. In addition, the new publishing pop songs have a higher degree of popularity among music streaming listeners. The finding of this study may support the music streaming to better understand their listeners, and to provide a guide reference for the traditional music industry.

**Keywords:** Chinese pop music, Trend analysis, Big data, Music streaming, KKBOX

過去判斷歌曲是否暢銷的方式，不外乎是藉由唱片銷售數字統計以及抽樣調查。但隨著大數據時代的來臨、數位串流音樂的興起，決定一首歌曲是否為「金曲」的方式逐漸的在轉變，從影音社群（如YouTube、SoundCloud）、音樂串流媒體（如KKBOX、Spotify）以及網路電台（如Pandora）等音樂提供平台所產生的大量聽眾收聽資訊，已經改變了各大音樂排行榜判斷熱門音樂曲目 的演算法，甚至改變了音樂產業的行銷與營運方式(Fong, 2014)。如美國指標性的音樂雜誌告示牌，其排行榜改變過去只統計唱片銷售數字及廣播播放次數決定排行榜順序，他們開始加入網入社群網路與音樂串流媒體（如YouTube、Spotify）上龐大的使用者播放資料，改變排名的演算法，找到真正的「金曲」。

根據文化部 2012 台灣流行音樂產業報告(張崇仁, 2013)，台灣的唱片產業主要分為四個構面：唱片銷售、數位音樂、現場演出、音樂版權應用。其中，經營有聲出版業者經營數位音樂者所佔的比例為74%，其總產值達 17.76 億。在經營數位音樂者 89.19%有「Internet」相關業務，並且全數皆有與KKBOX合作。

為了解華語流行音樂之趨勢，本研究將採用KKBOX每日華語流行音樂前一百名排行榜，藉由網路爬蟲挖掘自2005年至2015年共340165筆音樂排行榜資料，並篩選出一百首最熱門的歌曲進行分析，探索華語流行音樂發展。

本研究期望能協助數位串流音樂服務，能更了解其聽眾喜好，並提供傳統音樂產業營運之參考。

**1 文獻綜述**

1.1大數據

大數據（Big data）指的是以現有的科技難以 處理的大量資料,資料的大小並沒有被定義,而是依照當時的科技能力而定(Manyika et al., 2011)；大數據的資料類型可分結構化資料（即關聯式資料庫能夠容易處理之資料類型，如數值、字元字串、布林值等）與非結構化資料（即關聯式資料庫難以直接處理之資料型態，如網頁、文件、多媒體等）。大數據具有4V的特徵(胡世忠, 2013; Laney, 2001)，分別為：Volume(大量)，意指資料的數量，而大數據的資料量通常是以現有的科技能力難以處理的數量，會隨著科技的演進不斷的增加；Velocity(快速)，意指資料產生與更新的速度 ﷽﷽﷽﷽﷽﷽﷽﷽，例如購物網站不斷快速產生的交易資料或者社群網站上會員的活動紀錄等，每分每秒都在快速地增加與更新；Variety(多樣性)，意指資料的內容與結構有的豐富的多樣性，除了結構化的數據資料，也存在著非結構化的圖像、聲音、影片及社群網站上的推文內容等；Veracity(真實性)，意指資料本身可靠程度、可信程度，若資料本身的品質就已經存在問題，分析出來的結果必然會有問題。

由於大數據有著以上特性，配合資料分析的技術，能夠從原本看似沒什麼用處的數據資料分析出有用的資訊，並且進而轉換成商業智慧或者洞悉出知識,協助企業與科學找出問題背後的本質，了解目前的狀況,進而分析未來的趨勢。

### 大數據最有名的案例非沃爾瑪的「啤酒與尿布」莫屬，利用消費者的消費資訊並利用關聯分析，成功找出購買尿布的男性客人，通常也會順手帶一手啤酒回家。沃爾瑪藉由這些消費者購買商品看似沒有任何關聯的資料中，藉由大數據的 分析方法挖掘出成功提昇企業競爭力的重要知識(WIRED, 2003)。

### 大數據的應用範圍並不侷限於某種產業，舉凡金融、醫療、航空、服務業、 半導體生產製程等各式各樣的行業中都有應用成功的案例，對於音樂產業也是如 此，如美國《告示牌》雜誌中的告示牌百強排行榜（Billboard Hot 100）是美國指標性的音樂單曲排行榜，統計週期是由每週一至週日，過去利用銷售數以及廣播數來決定誰能登上排行榜。但是在 2013 年三月，它將 YouTube 的播放率加入了演算法中，使其重視這點的歌曲就是「哈林搖」。「大賣才能成為金曲」這個觀念已經是個過時的觀念，「江南 Style」以及 「Call me maybe」也都是同樣的例子。因此，告示牌除了將 YouTube 播放次數 加進他的演算法外，也加入了其他主流串流媒體上大量的數據，進而改善了他的 排行榜。現在的流行音樂要成為金曲不能只從銷售的數字決定，反而聽眾開始掌 握這個決定權，音樂產業則必須開始重視並且應用這些社群間分享的大數據(Times, 2013)。

### 1.2流行音樂

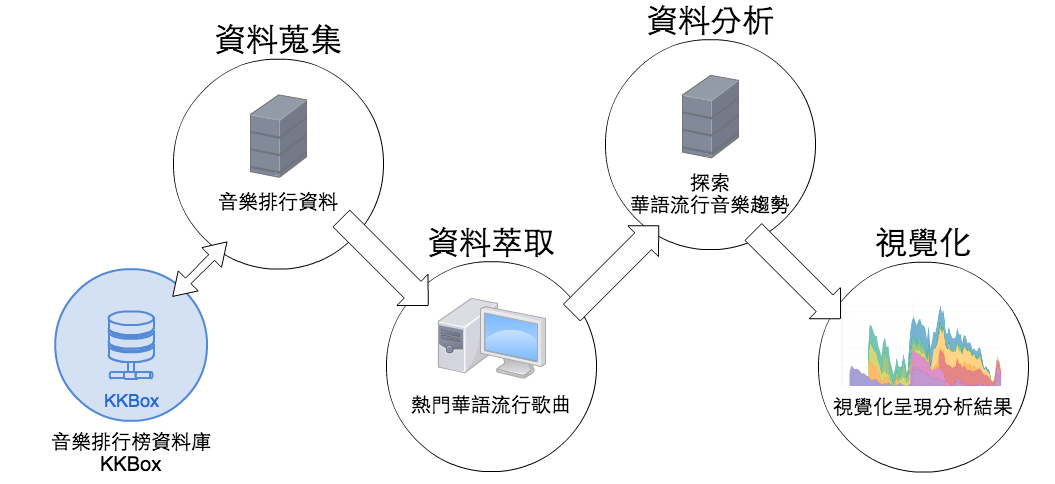
### 據(Wikipedia, 2015)描述,華語流行音樂泛指流行音樂中以中文為主要演唱語言的流行音樂,為台灣音樂產業中相較於台語、粵語，較為主流的流行音樂,主要流行於亞洲使用華語的地區，如台灣、香港、新加坡、馬來西亞、中國大陸等地區,甚至在日本韓國也有其市場。

### 早期華語流行音樂主要是以唱片產業為主，80、90年代唱片產業發展達到了最高峰，各國唱片產業紛紛加入這場戰局。但隨著電腦技術以及網路的發展，盜版唱片與P2P重重的打擊了傳統音樂產業，唱片銷售量開始一蹶不振，面臨轉型(施韻茹, 李天鐸, & 郭良文, 2004)。

### 由於網際網路的發展，數位串流音樂開始興起根據訪談，大多數的音樂產業業者認為數位音樂將會是音樂產業市場的核心業務。在數位音樂的營收結構中，串流音樂收聽佔總產值 40%，顯示串流音樂在數位音樂中之重要性。

## **2 研究方法**

相較於實體唱片銷售數字，音樂排行榜能夠較直接反應聽眾對於流行音樂的喜好，具有較高的指標性(曾湘雲, 2003)。為了了解聽眾對於華語流行音樂之喜好，本研究藉由Python撰寫網路爬蟲程式，蒐集KKBOX每日華語流行音樂前一百名排行榜，藉由網路爬蟲挖掘自2005年至2015年共340165筆音樂排行榜資料，萃取出前100首最熱門的曲目，進而分析華語流行音樂之發展趨勢，並且視覺化呈現分析之結果。圖為本研究之架構。



2.1資料蒐集

本研究資料蒐集來源為目前台灣具指標性的線上音樂串流平台 KKBOX 所提供的華語流行音樂排行榜及點播相關資料。KKBOX 提供之音樂熱門排行榜資料，會於每日統計正式會員播放的資訊，並進行統計分析，排名出每日的熱門音樂排行榜，其排行榜名次無法藉由買賣、關說及個人意見影響與干涉，可確保其公正性。

KKBOX提供從2005年起至今的音樂排行榜，依照音樂頻道做分類，分為綜合新歌、華語、西洋、日語、韓語、台語、粵語、嘻哈 R&B、搖滾、電子、古典、爵士以及世界心靈共十三類，依統計週期分為日排行榜、週排行榜及月排行榜，統計並分析出最熱門的前一百名單曲以及專輯。

由於排行榜資料量龐大，難以藉由人工手動蒐集，因此本研究利用Python撰寫網路爬蟲，蒐集KKBOX公開之排行榜自2005年起至2015年，每日公布之前一百名熱門華語流行音樂排行，共340165筆華語流行音樂之排行榜資料，以供後續探索華語流行音樂趨勢之分析。

2.2資料萃取

為利用網路爬蟲取得之大量資料了解熱門華語流行音樂之發展趨勢，本研究除了利用排行榜探尋華語流行音樂趨勢之輪廓，並對這些資料進行篩選，剔除熱門程度不高之曲目，萃取出這十年間最熱門100首最熱門之曲目，作為較深入的對象。

2.3資料分析

為分析這些大量的資料，本研究使用R語言進行資料的篩選與分析。R語言是一種跨平台開放原始碼之程式語言，主要用於統計分析、資料探勘以及資料圖表繪製。其優點為已內建多種統計學函式以及數字分析的功能，並且擁有廣大之社群提供大量且多樣的套件（Packages，為用戶撰寫之功能函式）。R語言也具有強大的繪圖功能，有助於探索與發現從資料數值上面難以直接發現的線索。

2.4視覺化

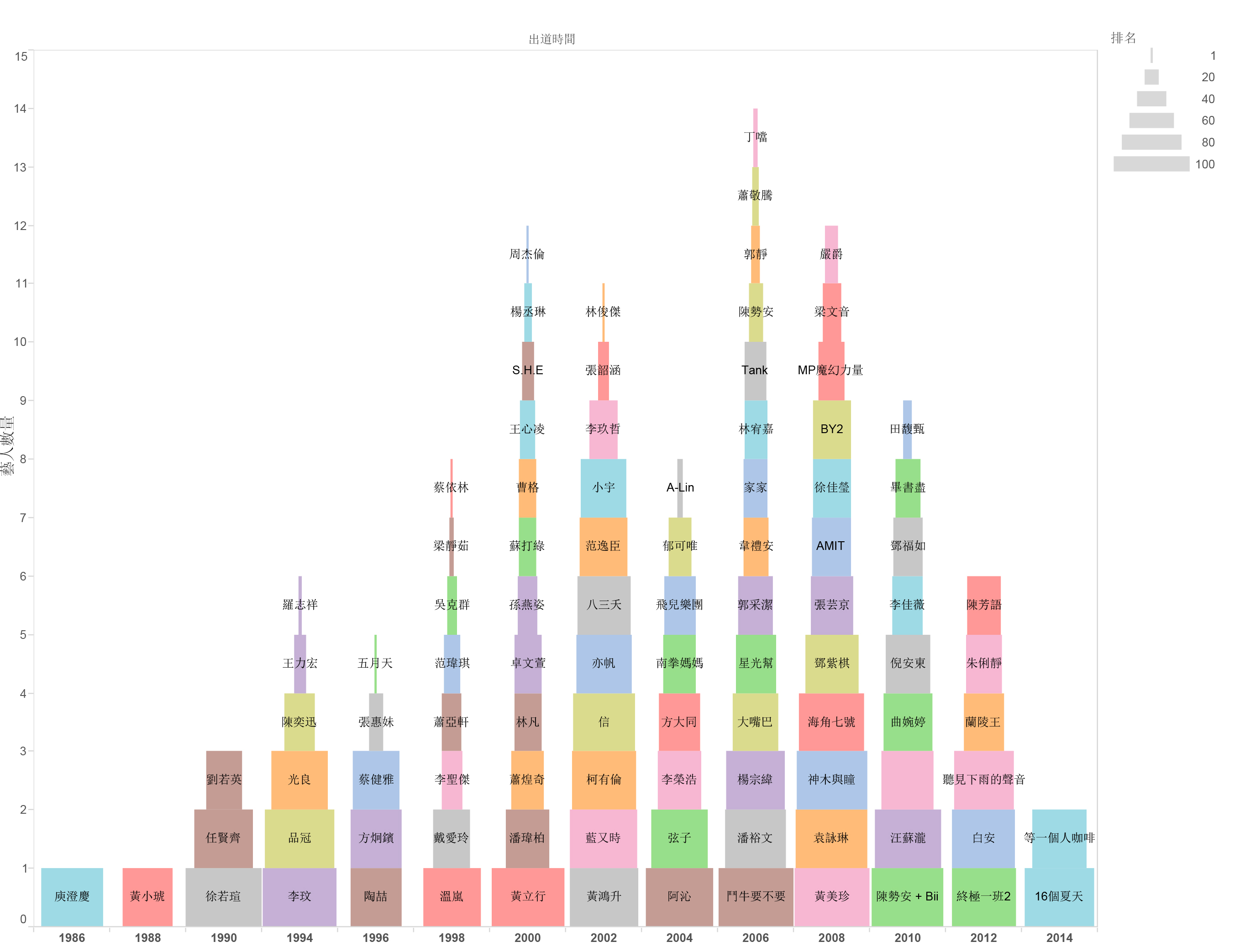
資料視覺化（Data Visualization），為了更清楚呈現資料所代表之意義，除了進行統計分析外，本研究利用R語言本身之繪圖功能加上商業視覺化分析軟體Tableau，將難以直接解讀之資料分析結果，以視覺化圖形精確呈現，有助於研究探索華語流行音樂之趨勢以及現象。

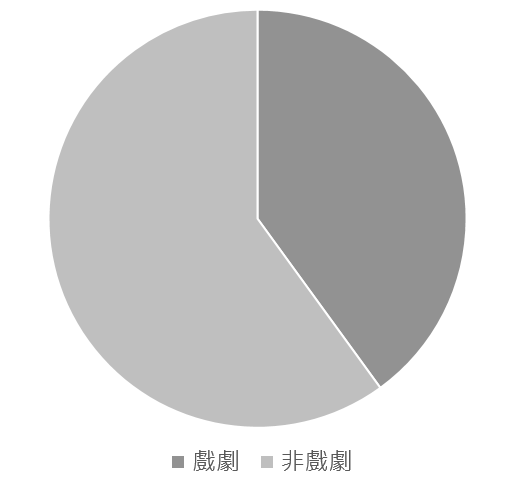
## **3 研究發現**

本研究主要針對流行音樂結構概況進行分析，接著從中探索流行音樂發展現象與趨勢，並配合視覺化圖表觀察背後可能代表之含義，以下將分作兩點說明，分別是流行音樂概況以及影視戲劇對於流行音樂之影響。

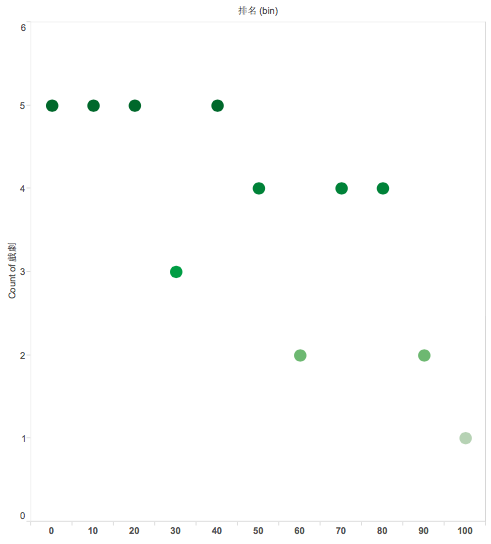
3.1流行音樂概況

由於華人分佈於世界各地，華語流行音樂可能來自於不同國家，為了解熱門華語流行音樂的創作來源，本研究利用萃取出之前一百名熱門華語歌曲演唱之藝人，結合地圖資訊，描繪出地區分布圖。圖中顯示，約69%之熱門藝人來自於國籍為台灣之藝人，其次10%來自於國籍為美國之藝人以及8%來自於馬來西亞之藝人。

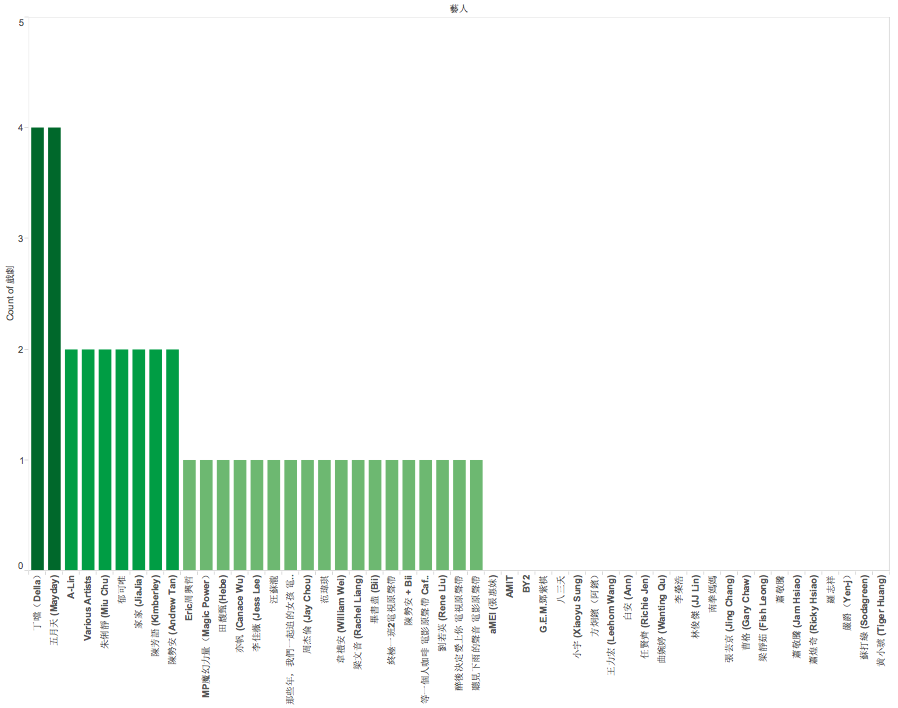
藝人出道之年份與時間長度可能會影響聽眾對於藝人之認識程度，進而影響對於藝人演唱歌曲之喜好，本研究整理了每個時期出道藝人，比較哪個時期的藝人為現今華語流行音樂圈中較為活躍，圖為前一百名藝人出道時間之分佈情形，圖形越細的藝人代表名次越頂尖。從此圖中可發現活躍的藝人大多集中於1998年至2008年這段時間，正值網際網路興起，推測在這段時間活躍之藝人較懂得使用網際網路行銷，因此曝光程度較高，線上串流之聽眾較多；於1998與2000年出道之藝人較為熱門；另外2006年出道之藝人數量最多，可能原因為當時音樂選秀節目盛行，產生藝人之數量較多；這些藝人中，有部分屬於影視戲劇之原聲帶，推測影視戲劇對於歌曲熱門程度可能具有影響力。

藝人的出道的形式可能也會影響聽眾之喜好，因此本研究將藝人的形式分為男藝人、女藝人以及團體藝人，並且進行比較。圖中，每個區塊大小代表各類型藝人所站之比例，以個人名義出道之藝人佔有相當大的百分比，以樂團形式或者男女團體形式之藝人佔有比例相當小，可以推測以團體形式在台灣出道可能較不合華語聽眾之喜好；在比較過程中，發現以影視原聲帶發片之形式也佔有不小比例，因此也加入此比較中，並且在下一個階段針對影視戲劇相關曲目進行分析。

3.2影視戲劇對於流行音樂之影響

由前面幾個觀察可以發現，影視戲劇對於歌曲的熱門程度具有一定程度的影響力，本研究針對影視戲劇相關曲目進行探索，圖為前一百名影視戲劇相關歌曲所佔之比例，影視戲劇相關曲目佔40%，其餘60%為非影視戲劇相關曲目；接著，本研究將影視戲劇相關曲目名次分佈繪製成圖形，以10名為一個範圍，顯示該範圍中屬於影視戲劇相關的曲目數量。從分佈圖中可以發現，影視戲劇相關曲目分佈於較前面的名次，可以推測影視戲劇可以影響流行音樂曲目的熱門程度。

另外，本研究統計了影視相關曲目與藝人之間的關係，哪些藝人較常利用影視戲劇為行銷之方式，如圖顯示，以丁噹與五月天演唱，並且配合影視戲劇的曲目都為四首，所佔之比例最高。



## **4 結論與討論**

本研究透過蒐集數位串流音樂平台KKBOX提供之華語流行音樂每日前100名排行榜，探索華語流行音樂發展趨勢，發現熱門華語流行音樂主要發源地為台灣（69%），其次為美國（10%）以及馬來西亞（8%）；熱門華語流行音樂演唱之藝人出道時間主要分布於1998年至2008年，推測是由於網路發展，較懂得利用網路行銷增加曝光度；2006年為台灣音樂選秀節目興起，使得此年出道之藝人數量達到一個高峰；統計藝人類型比例可發現，以個人身份出道之藝人較符合聽眾之口味；另外以原聲帶發片之歌曲也佔有一定的比例，因此本研究接者探索這個現象。

統計前一百名屬於影視戲劇相關之曲目，佔全部比例40%，將近一半，並且主要分布於較前面的名次，表現突出，可推測影視戲劇可能對於歌曲熱門程度具有相當的影響力；本研究統計哪些藝人所演唱之歌曲為影視戲劇相關歌曲之大宗，發現佔最大比例之藝人為丁噹與五月天，可推測他們已分線此現象甚至懂得利用此現象行銷其歌曲以及增加曝光度。

本研究為一個概括性的統計，統計2005年至2015年所有之曲目並進行篩選分析，了解華語流行音樂趨勢之輪廓，因此並未依照每個時期進行分析比較，為了解更精確之趨勢，未來可以將所有上過榜之曲目與藝人納入統計分析，並且依照每個時期可能發生之事件進行對照，以了解背後所代表真正的意義。最後，本研究希望可以協助數位串流音樂服務，能更了解其聽眾對於華語流行音樂之喜好，並提供傳統音樂產業營運之參考。

**參 考 文 獻**

施韻茹, 李天鐸, & 郭良文. (2004). *由節拍旋律到娛樂商品: 台灣流行音樂產業產銷結構轉變研究.*

胡世忠. (2013). 雲端時代的殺手級應用：Big Data海量資料分析. *天下雜誌*.

張崇仁. (2013). *101 年流行音樂產業調查報告*

*TAIWAN POP MUSIC INDUSTRY SURVEY 2012*: 文化部影視及流行音樂產業局.

曾湘雲. (2003). *檢視台灣流行音樂市場結構與產品多樣性之關聯性: 從歌曲內容及音樂產製面談起.*

Fong, D. (2014). How big data can change the music industry. from <http://venturebeat.com/2014/12/18/how-big-data-can-change-the-music-industry/>

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note, 6*.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., & Roxburgh, C. (2011). Big data: The next frontier for innovation, competition, and productivity.

Times, T. N. Y. (2013). What’s Billboard’s No. 1? Now YouTube Has a Say. from <http://cn.nytimes.com/business/20130316/c16billboard/en-us/>

Wikipedia. (2015). 華語流行音樂. from <http://zh.wikipedia.org/wiki/>

WIRED. (2003). 【WIRED特刊】Wal-Mart轉型科技企業的關鍵報告：看它用Big Data算出你的「消費基因」！. from <http://www.wired.tw/posts/bdma_2_4_walmart_social_genome>

1. 基金專案：科技部專題研究計畫案，編號MOST-103-2410-H-158-010, 實踐大學專題研究計畫案，編號USC 103-05-04002

   通信作者：廖偉帆，實踐大學資訊科技與管理學系，研究生，e-mail：p988744@gmail.com。 [↑](#footnote-ref-1)