

Eyes on Singapore and Hong Kong: An Interactive Visual Application to Explore the Start-up Ecosystems

Chen Yuge, Fu Yu, Zhuo Yunying

Abstract—Blockchain, artificial intelligence, data science, fintech and internet-of-things are all buzzwords today in this new innovation era. The global start-up revolution will continue to evolve and grow with the emergence of new technologies and new business models. The underlying contribution to innovations lies in having a strong and robust start-up ecosystem. Many developing or developed cities have already recognised the significance of entrepreneurship activities. However, the success of the ecosystem is affected by multiple factors, signifying the need to enhance start-up ecosystem management through more data-driven comparisons. This research project aims to take an interactive and exploratory approach to study the differences between the start-up ecosystems in Hong Kong and Singapore in terms of funding support and start-up profiles. To achieve this, our team designed an interactive visual application utilizing data from Crunchbase. These visualizations mainly make use of R shiny framework and numerous R packages to display the differences. Through this data visualization application, our team hopes to assist entrepreneurs, policy makers and investors in decision-making through better understanding the state of start-ups in different industries across the years.

Index Terms—Visual Analytics, Start-up Ecosystem, Funding, Investment, Singapore, Hong Kong

1 INTRODUCTION

With the arrival of the Third Wave of Innovation, increasing number of global start-ups are disrupting established industries by revolutionizing new technology frontiers. The start-up landscape also begins to shift to Asia, coupled with the exponential growth in start-ups and investments in the region. Establishing and building a robust start-up ecosystem becomes even more crucial for many developing and developed cities in Asia today. In a bid to improve on the existing ecosystem, comparisons with other start-up ecosystems help to draw essential insights and lessons especially if the economies are largely similar.

Till date, there exists limited data-driven comparisons between start-up ecosystems. One of the more comprehensive sources would be the annual Global Start-up Ecosystem Report published by Startup Genome. The ecosystem analysis had factored in resources in terms of talent, funding, founder and organization, networks in terms of local connectedness as well as performance in the area of number of start-ups generated. However, the comparisons on start-up ecosystems tend to be confined to static charts with simple and highly aggregated start-up ecosystem attributes. This is mainly due to a general lack of consolidation and standardization of the information on start-ups and investors across countries as the nature of start-up ecosystem is volatile and ever-changing. Furthermore, the choice of static charts has restricted the usage and interactivity of visualizations.

This paper aims to provide a client-based interactive and dynamic web visualization tool for start-up ecosystem comparison on the disaggregated level. Data are sourced from CrunchBase, a platform that provides information on both private and public start-ups. Our aim is to understand the differences between the two similar city-states (namely Hong Kong and Singapore) in the aspects of information on start-up companies and funding organizations within the ecosystem. Other players including government funding and crowdfunding platforms are neglected due to the lack of consolidated historical data.

2 MOTIVATION AND OBJECTIVES

As there is a general lack of effective and user-friendly visualization for discovering country-specific differences between the start-up ecosystems, the main motivation behind this project is to create a centralized, dynamic and interactive dashboard for quantitative comparisons on various aspects of start-ups and funding organizations in different countries. We believe that through the usage of interactive data visualization techniques, we will enable end-users to explore the data via numerous filtering and slicing features. Based on User-Centric Dashboard Design Guide, this dashboard

takes a broad, strategic, customizable, drillable and exploratory approach and is targeted at potential entrepreneurs, policy makers and investors. Specifically, it attempts to support the following analysis requirements.

- 1) Enable potential entrepreneurs to understand the growing and declining industries, investors' profile as well as pinpoint the top funded start-ups
- 2) Help policy makers to identify the potential profitable and leading industries and dedicate more resources in specific industries
- 3) Assist investors in identifying the difference between Singapore and Hong Kong's start-up industries and strategizing future investments in these two regions

3 RELEVANT WORKS ON START-UP ECOSYSTEM

There has been work done on the visualization of start-up ecosystems through static charts and figures. The figure below shows the Global Start-up Ecosystem Report 2018 published by Startup Genome and it illustrates the country-specific aggregated attributes on different ecosystem indices including local connectedness and ecosystem demographics. However, the figure contains relatively high level of details and presents a one-size-fits-all snapshot of prescriptive data. The highly static and non-customizable features require users to source for figures on countries individually which defeats the purpose of comparison.

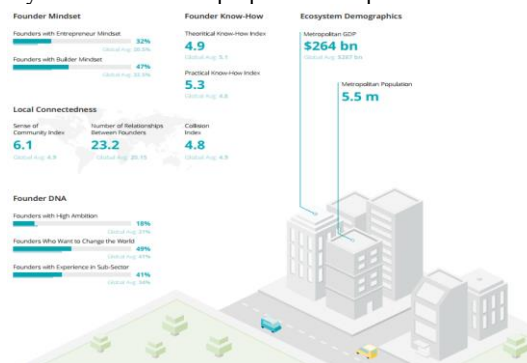


Figure 1: Start-up Ecosystem Visualization

Another figure (Figure 2) published by Startup Genome in its Global Start-up Ecosystem Report 2017 shows a heat map that display time-series

data for multiple cities across categories such as valuation of exits above \$100 million, valuation of exits above \$1 billion and start-up valuation above \$1 billion. The colour intensity allows the users to highlight pattern of data in quantity and identify outliers easily. In addition, low ink-ratio is used in this heatmap to facilitate easy reference across dimensions. However, the start-up ecosystems comparisons are done on countries with drastic cultures, economy characterises and region-specific attributes. Background of cities are not taken into consideration while users are unable to drill down into specific industries.

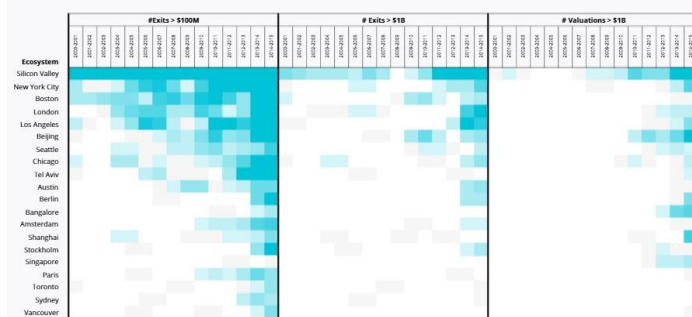


Figure 2: Heatmap on Different Start-up Ecosystems

Figure 3 published by Tech in Asia indicates the attempt to present top-funded start-ups in Singapore. While the information presented are clear and succinct, pure listing by start-ups by funding amount lacks interactivity and drillable details. The visualization has exceeded the boundaries of a single page while users must click on specific icon to view further details in separate screen. This has violated the good dashboard practices as mentioned by Stephen Few and we aim to correct this in our visualization.

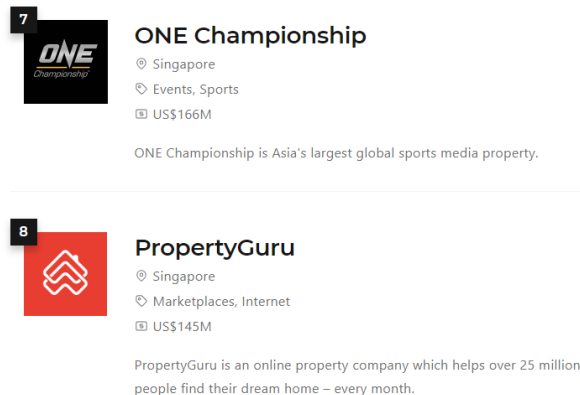


Figure 3: Listing of Top Funded Start-ups

4 THE APPLICATION

4.1 Data Preparation

The datasets for this project contain detail information of start-up companies, funding information break down by funding stages, and investor details and were retrieved from CrunchBase. There are two limitations in our dataset. Firstly, despite the large number of attributes provided in dataset, the attributes which are qualified to be used for analysis is limited, due to the high sparsity of data caused by missing values. Secondly, the exact information of investors' past investment is un-disclosed due to privacy issues. After prudent trade-off between the dataset quality and the desired start-up analysis outcome, we decided to focus our visualization on start-ups' industries and funding amount analysis.

One of the challenges we faced during data cleaning is the start-ups categorization. In the original dataset, there are multiple categories for each start-up and they are non-hierarchical, unstructured and sometimes repetitive. This is attributed to non-standardization of industries in CrunchBase due to multiple data sources and nature of start-up industries being new and emerging. To handle this, we firstly consolidated a

reference list of categories with two hierarchies "sector" and "Industry", and then used python script to filter out "common" categories between the reference list and dataset. Lastly, we manually looked through start-ups belonging to more than one category and choose the right category based on our business knowledge. In addition, we joined start-ups information data with investments data to link start-ups with their funding for different stages. We also performed geocoding for investors country of origin, and manually filled in 200 missing values by searching the personal information of investors.

Based on information visualization reference model, data transformations are done according to the requirement of specific data visualization charts we have chosen. This is to ensure visual structures are presented with the right visual mapping from the data tables.

4.2 System Architecture

The project was developed using R language, due to its advantage in data manipulation, availability of extensive open-sourced mature packages for interactive plots. We used shiny framework for interactive web supplication and "shinydashboard" for web design. The advantage of shiny is that it allows us to build customized web apps directly from R, and incorporate tableau public workbook, CSS, htmlwidgets and JavaScript easily when required. The application is deployed, and the data was decided to be stored as plain CSV format for ease of access.

5 INTERFACE DESIGN

5.1 User Design Considerations

The application is designed for different users including entrepreneurs, investors and policy makers to navigate according to their preferences. The intention of the entire dashboard is to allow users to explore and thus user-driven interface is adopted.

The dashboard's usability is further enhanced through applying principles from Robert E. Roth's visual analytics objectives primitives and operant primitives for user Interface Design. A combination of identification, comparison, ranking and delineation was used in different visualizations. In addition, users can pan and zoom in some of the visualizations to drill down to details or visualize tooltip. Various versions of filters including year slider, indicator filter, map background adjustment and visualization types filter are also incorporated. To enhance aesthetics and minimalistic design, blue and red colour palettes are specified for all charts on Singapore and Hong Kong respectively.

We have also decided to adopt Ben Schneiderman's mantra – "Overview first, zoom and filter, then details-on-demand" in designing our overall dashboard. The shiny dashboard follows a 2-step approach where users first have a macro overview of the start-up ecosystem of Hong Kong and Singapore under the Summary Page and then micro perspectives under other tabs on understanding various key attributes depending on the user type.

5.2 Selected Visualizations

5.2.1 Summary Page

A summary page (Figure 4) will be shown when users open the web application. It shows a highly aggregated overview of Singapore and Hong Kong start-up ecosystem. Under this page, total number of active and inactive start-ups and time-series start-up formation trend are shown through summary statistics, line chart and stacked bar charts.

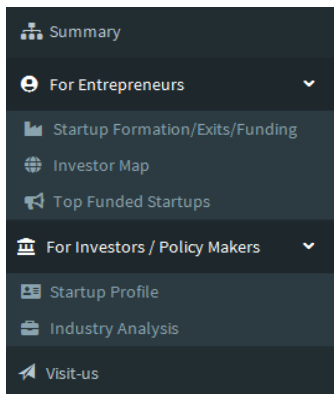


Figure 4: Side Bar on Rshiny Dashboard

5.2.2 For Entrepreneurs

(A) Heatmap: Start-up Formation/Exits/Funding

The heatmaps (Figure 5 and 6) allow users to compare the trend of formation, exit and funding of start-ups in different industries across the years. Users can choose to view the start-ups formed/exited and funding amount by industry in both Singapore and Hong Kong from 2000 to 2018. Users can have an idea of which industries are getting more popular among other entrepreneurs and investors and consider which industry to enter based on this information.

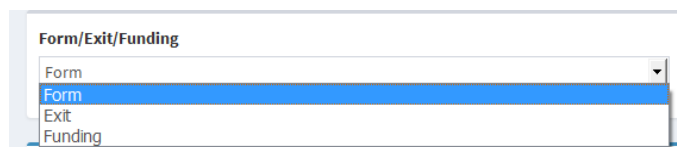


Figure 5: Filter Based on Different Indicators



Figure 6: Heatmap on Singapore and Hong Kong

(B) Bubble Map: Investor Map

Investor origins are displayed on world maps to display spatial patterns (Figure 7). Different bubble colours and sizes enable users to identify differences in number of investments within the economy and between economies as bubble size represents investment quantity while colour represents percentile. Map background customization function allows users to view the bubbles more clearly in case the bubble colours clash with the map background. A miniature map function is also incorporated for smooth navigation on top of drag feature on the map. For easier comparisons, the maps are put side by side and synchronised in such a way that they can zoom in and out together and when users hover to place on one map, the same place on the other map is highlighted in a red circle.

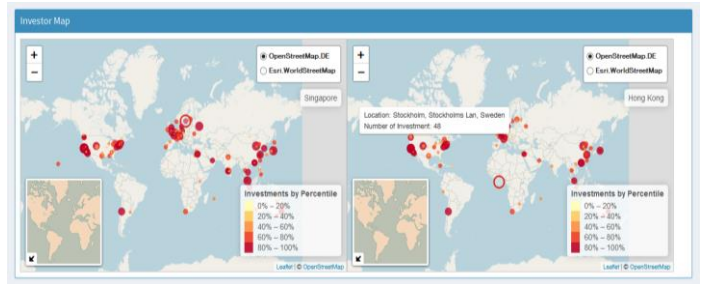


Figure 7: Interactive Bubble Map on Investors

(C) Bar Charts and Data Table: Top Funded Start-ups

We ranked the start-ups based on their total funding amount. Users can choose company headquarter location, industry and top N to filter the data shown (Figure 8). Besides showing the total funding amount, we also include the funding raised at each funding stage in the tooltip. Users can read more about a company in the data table which shows the company description, founding year and total funding amount.

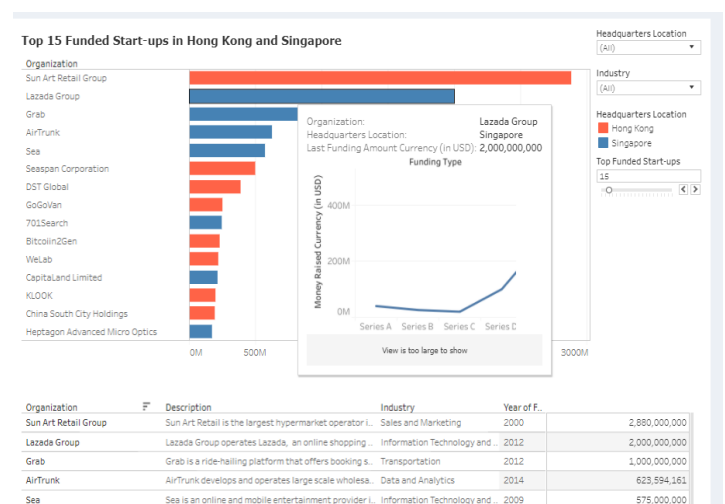


Figure 8: Tableau Bar Charts and Data Table

5.2.3 For Investors or Policy Makers

(A) Cleveland Dot Plot: Start-up Profile

Cleveland dot plot (Figure 9) is used because it is effective in comparing data of two parties on a single chart and easily communicate the differences between values. Moreover, unlike bar chart, this dot plot reduces redundancy and clutter and maximize the data-ink ratio. A time slider (Figure 10) is also implemented for users to see the changes across the years from 2000 to 2018. Investors and policy makers can make better decisions based on comparison on Cleveland dot plot. For example, investors can decide which industry they should invest more if they are interested in Singapore market while policy makers can appreciate the rising or declining industries.

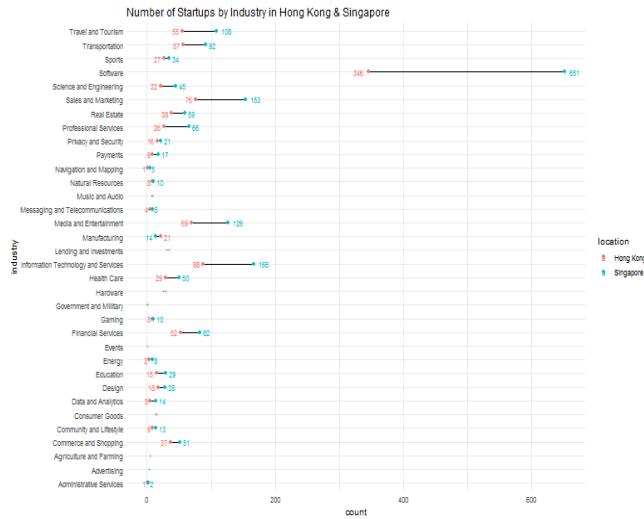


Figure 9: Cleveland Dot Plot based on Different Industries

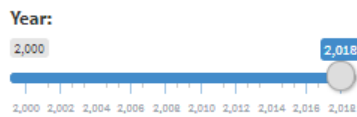


Figure 10: Time Slider from Year 2000 to Year 2018

(B) Violin Plot and Bar Chart: Industry Analysis

Under this session, users can drill down into details to compare the performance indicators between Singapore and Hong Kong start-ups. There are two main categories of indicators. The company-specific indicators include company age and CrunchBase Ranking while other marketplace indicators include SimilarWeb indicators and BuiltWith Active Technology Count. SilimarWeb is an online web intelligence platform while BuiltWith closely monitor the technologies utilised by different start-ups.

As displaying all the 34 industries in one chart results in difficulty for visualization, we grouped industries into different sectors. Users can view start-up performances at both sector and industry level within one single page (Figure 11). By incorporating filter functions, users can select a sector to focus on while information for industries under the selected sector will be shown in second row.

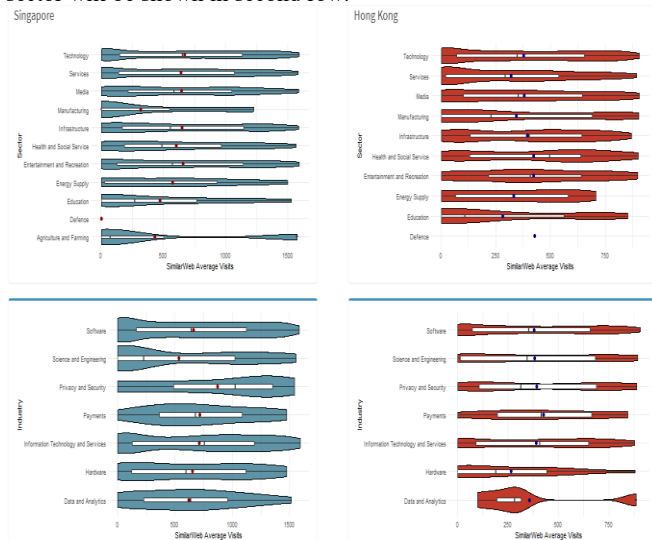


Figure 11: Violin Plot based on Industries and Sectors

In addition, users can also choose to filter violin charts for distribution analysis or bar charts which show summaries of each sector and industry performance under each indicator. A violin chart (Figure 12) combines

box-plot and distribution curve. It provides users with information on distribution and density.

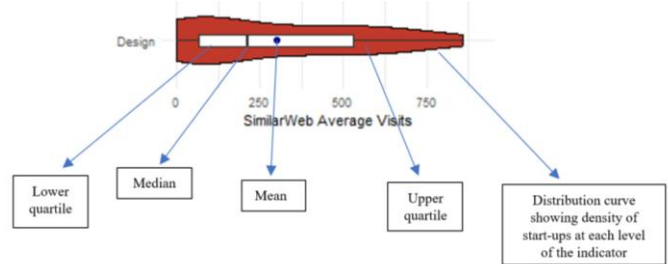


Figure 12: Detailed Explanation on Violin Plot

Bar charts showing total values are implemented only for indicators “SimilarWeb Average Visits”, “SimilarWeb Visit Durations” and “BuiltWith Active Tech Count” while error bar charts are used for “Age (Days)”, “CrunchBase Rank” and “SimilarWeb Global Traffic Rank”. Bar chart is less suitable for indicator Age (Days) as a higher total age of the sector could be attributed to the higher number of start-ups in that sector, resulting in misleading information. As indicators CrunchBase Rank and SimilarWeb Global Traffic Rank contains ordinal values, bar chart with sum values is also inappropriate as compared to error-bar charts. Error-bar charts (Figure 13) are more effective as the bars show mean values and the lines are useful in displaying standard error, which reflects how precise mean values are as measurements of performance.

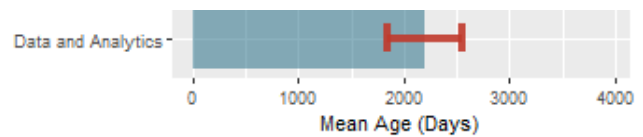


Figure 13: Detailed Explanation on Error Bar Chart

6 CASE STUDY: SINGAPORE AND HONG KONG

For the purpose of this case study, we shall look at the two Asian Tigers, Hong Kong and Singapore. Hong Kong and Singapore are two high-growth metropolitan city-states in Asia that share many characteristics in common in terms of GDP per capita, population density, high foreign direct investment, robust business infrastructure, limited natural resources etc. Furthermore, for both countries, the relevant government bodies have invested in considerable amount of resources to boost and build up the startup ecosystems. For instance, Hong Kong has set up a HK\$2 billion Innovation and Technology Venture Fund for co-investment in local innovation and technology start-ups. On the other hand, Singapore government has been pushing for various initiatives such as Block 71, Smart Nation, SkillsFuture to develop dedicated start-up clusters, good technological infrastructure and talent pool. Thus, building a robust start-up ecosystem aligns with the government and industries’ interests for both countries. To provide a realistic view of comparison, start-ups founded from 2000 after the dot-com boom are considered.

6.1.1 Results and Insights for Entrepreneurs

Based on start-up formation/exit/funding heatmaps, software industries are getting popular among entrepreneurs in both Singapore and Hong Kong while popularity of media and entertainment industry and information technology and services industries were once peak in 2013 and 2014 but dropped in recent years. Start-up exit numbers in both economies are almost negligible. Nevertheless, entrepreneurs should pay attention to Singapore information technology and services industry because exit-funding ratio in this industry is relatively higher. As for funding, travel and tourism, software, science and engineering and information technology and services industries in Singapore get more funding than those in Hong Kong while Hong Kong’s financial services industry gets more funding compared to Singapore. Hence, entrepreneur should consider Singapore industries such as travel and tourism, and

science and engineering as they are less competitive and yet more attractive to investors.

As seen on investor maps, both Singapore and Hong Kong are attractive to investors from the United States, Europe and China. While the patterns in the United States and China are similar for the two economies, Singapore has more diverse European investors. Moreover, investors from Moscow invest much more in Singapore than in Hong Kong. Singapore entrepreneurs can thus consider investors from Moscow investors or other European countries who have not invested in Hong Kong. With less competition, the chance of securing funding could be increased. However, as the funding stages of the funded amounts are not disclosed fully, the success rate in different funding rounds might still differ across investors from different regions. Entrepreneurs should exercise cautions in pitching to the right investors.

6.1.2 Results and Insights for Investors/Policy Makers

From the Summary Tab, we can visualize the trends in the startup formation across the years very clearly. For both Hong Kong and Singapore, the changes in the number of start-ups formed per year from 2000 to 2018 follow a similar pattern. From 2000 to 2008, there is a relatively small number of new start-ups formed each year, ranging from 10 to 100. Subsequently, there is an exponential growth in the number of start-up formations from 2008 to 2014. From the year 2014 onwards, the number began to decrease consistently across the years. The drastic declining rate of home-grown entrepreneurship raises huge concern for both Singapore and Hong Kong. In spite of strong government support, the rapid decrease in start-up formation signifies a greater challenge for both countries. There is a lack of risk-taking and entrepreneur-ready culture and mindset among the local younger generations today. The problem is much more serious in Hong Kong, considering the lower number of start-ups than Singapore. The relevant government agencies should find new ways to cultivate failure and entrepreneur mindset for the next generations.

As observed from Cleveland dot plot, over the years, travel and tourism, software, sales and marketing, media and entertainment and information technology and services industries are experiencing great increase in number of start-ups. As of October 2018, Singapore has advantages in travel and tourism, software, sales and marketing, media and entertainment and information technology and services in terms of number of start-ups compared to Hong Kong, though these industries are also leading in Hong Kong. Investors should consider Singapore start-ups for these industries as the ecosystem is more mature in Singapore.

Singapore's better start-up ecosystem than Hong Kong could be resulted from aggressive initiatives in recent years by drawing lessons from Israeli entrepreneurship model. Relevant government agencies such as Media Development Authority, Infocomm Development Authority and SPRING Singapore has launched various programmes including University Innovation Fund, Proof-of-concept grants, Technology Incubation Scheme and Early Stage Venture Capital to encourage seed and early stage entrepreneurship funding. As these schemes have strong focus in technology-inclined business ideas, the schemes have been effective in sustaining the momentum in entrepreneurship activities in more knowledge-based industries. Unlike Singapore, Hong Kong's funding schemes are heavily tied to Research and Development as well as cumbersome reimbursement process. Singapore should execute policies to strengthen the comparative advantages in dominant industries and Hong Kong should consider further enhancement in its existing schemes or collaboration with Singapore to improve on these industries.

However, as mentioned earlier, the risk-taking culture in both Singapore and Hong Kong is still not strong as compared to western counterparts. There is a need to move beyond the dependency on government leadership and programmes.

7 CONCLUSION

The demonstration of the centralized and interactive dashboard underlines its ability to allow different targeted users to explore and analyse different attributes of similar start-up ecosystems based on information on start-ups and funding organizations. In the future, trend analysis and text analysis on the specific technologies utilized by the start-ups could be explored.

To obtain the full information of start-ups, we can consider scrapping information from start-ups' online webpage or form strategic partnership with local start-up government agency. For instance, Start-up SG network, the newly launched start-up platform would be able to provide up-to-date information. Visualization model could also be extended to accommodate other players of the start-up ecosystem such as government funding, support organizations (co-working spaces, crowd-funding platforms, accelerators), talent pool (universities courses, population diversity). Lastly, data richness and accuracy could be further enhanced by cross-referencing other start-up databases such as CBInsights, Dealroom.co, orb-intelligence.com, preqin.com and compass.co.

ACKNOWLEDGEMENTS

We would like to extend our gratitude to Dr Kam Tim Seong for this kind support and guidance throughout the IS428 coursework.

REREFERNCES

1. Broer, J. (2015, October 15). Hong Kong as startup hub for hardware startups. Retrieved November 25, 2018, from <http://jeffreibroer.com/2015/10/hong-kong-as-startup-hub-for-hardware-startups/>
2. Brown, S. (2015, January 8). Schneiderman's mantra - Coding the Architecture. Retrieved November 25, 2018, from <http://www.codingthearchitecture.com/2015/01/08/schneidermans-mantra.html>
3. Cleveland Dot Plots · UC Business Analytics R Programming Guide. (n.d.). Retrieved November 25, 2018, from <https://uc-r.github.io/cleveland-dot-plots>
4. Compass. (2016). Hong Kong Startup Ecosystem Report. Retrieved from <https://startupgenome.com/reports/Startup-Ecosystem-Report-Hong-Kong.pdf>
5. Error-Bar Chart. (n.d.). Retrieved November 25, 2018, from https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Error-Bar_Charts.pdf
6. Few, S. (2008). With Dashboards, Formatting and Layout Definitely Matter. Retrieved from http://www.perceptualedge.com/articles/Whitepapers/Formatting_and_Layout_Matter.pdf
7. Goh, M. (2018, April 11). Singapore and Hong Kong are winning over start-up accelerators. Retrieved November 25, 2018, from <https://www.cnn.com/2018/04/10/singapore-and-hong-kong-are-winning-over-start-up-accelerators.html>
8. National Institution of Standard and Technology. (2003, February 26). Violin Plot. Retrieved November 25, 2018, from <https://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/violplot.htm>
9. Reed, C. (2014, May 14). What is stopping the next generation of Singaporeans from becoming entrepreneurs? Retrieved November 25, 2018, from <https://sbr.com.sg/economy/commentary/what-stopping-next-generation-singaporeans-becoming-entrepreneurs>
10. Relander, B. (n.d.). Asia's Tech Start-Ups Are The Next Big Thing. Retrieved November 25, 2018, from <https://www.investopedia.com/articles/investing/021315/asias-tech-startups-are-next-big-thing.asp>
11. Roth, R. (2012). Cartographic Interaction Primitives: Framework and Synthesis. *The Cartographic Journal*, 49(4), 376-395.
12. Startup Genome. (2018). Global Startup Ecosystem 2018. Retrieved from <https://startupgenome.com/reports/2018/GSER-2018-v1.1.pdf>
13. Startup Genome. (2017). Global Startup Ecosystem 2017. Retrieved from <http://dl153wesras4r4.cloudfront.net/reports/GlobalStartupEcosystemReport2017.pdf>
14. Startup SG. (2018). Enterprise Singapore launches Startup SG Network, an online tech startup database, to enhance partnerships. Retrieved from https://www.gov.sg/~sgpcmedia/media_releases/enterprise-sg/press_release/P-20180917-2/attachment/MR02918_SSN%20Factsheet_2018%2009%202017.pdf

15. Tech in Asia - Connecting Asia's startup ecosystem. (2018, October 16). Retrieved November 25, 2018, from <https://www.techinasia.com/top-funded-startups-tech-companies-singapore>
16. Yeoh, F. (2016, May 6). Singapore's startup ecosystem: Have we arrived? Retrieved November 25, 2018, from <https://www.channelnewsasia.com/news/singapore/singapore-s-startup-ecosystem-have-we-arrived-8013936>