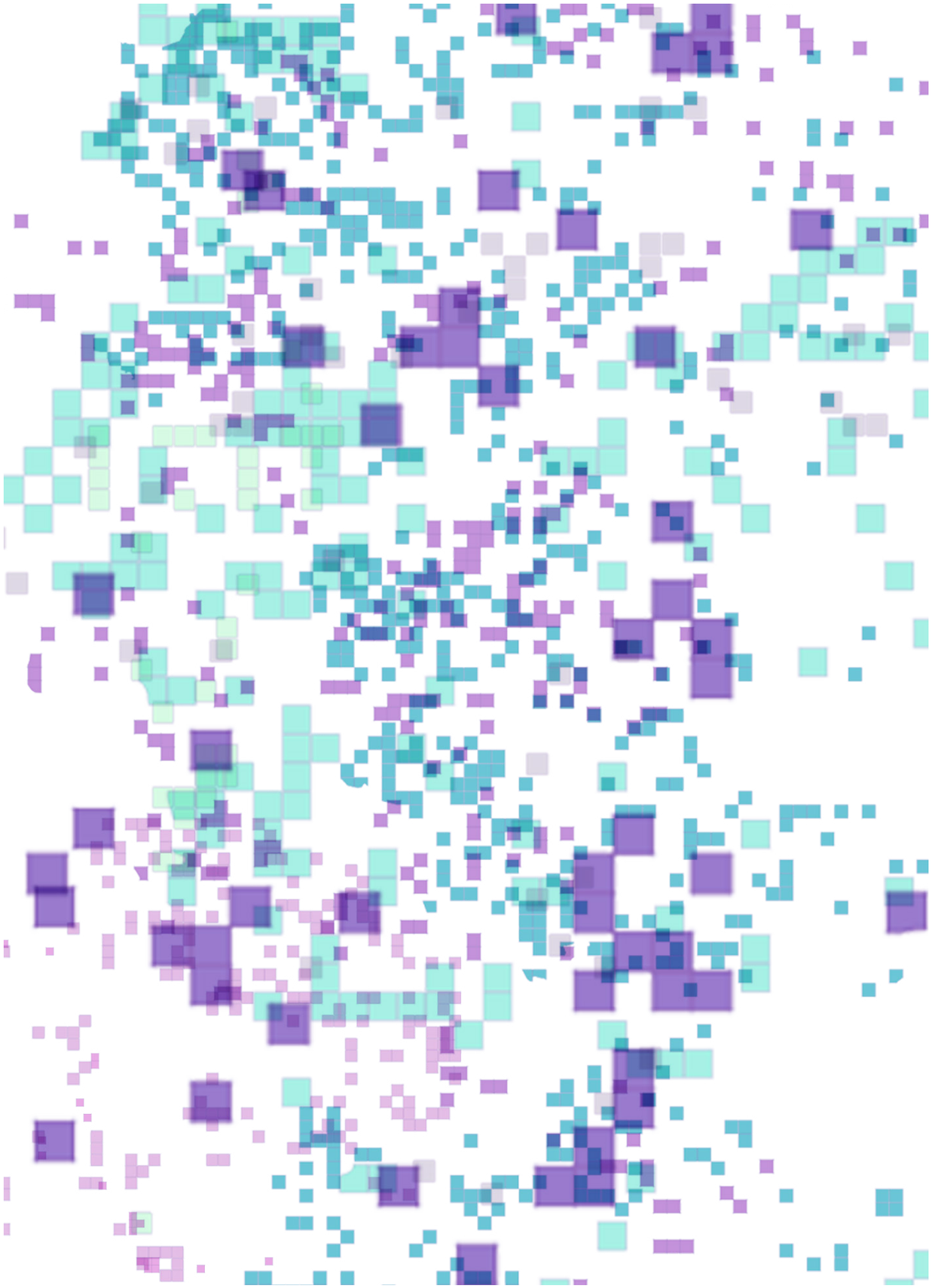


Potential Innovation District Study

Classification & Clustering Analysis & Planning & Design

A Personal Project

Period 08/2023-10/2023

The emergence of innovation districts has become a transformative force in urban spaces, driven by the influence of the creative class and the need for high-tech and creative industries. However, the existing literature highlights gaps in the classification and planning of innovation districts during their early stages.

Therefore, this project contributes to an enhanced understanding of potential innovation districts from the perspective of urban data and urban space. In the project, K-means clustering analysis is used to reveal seven distinct categories of innovation districts in Shenzhen, China. Asset-based strategies and design guidelines are further proposed for their planning and development.

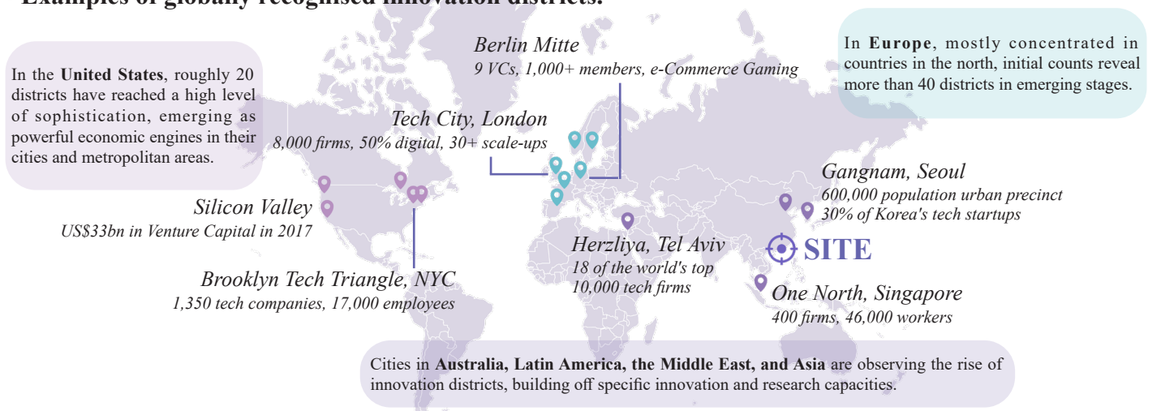
1. Introduction

1.1. Background

In 2002, "The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life", described a growing role of creativity in economy that involves a variety of fields, from engineering to theater, biotech to education, architecture to small business. And **3Ts**, technology, talent, and tolerance, are introduced as the new economic geography of creativity. Since then, many cities have been exploring land use policies and plans targeting the formation of innovative spaces.

In 2014, "The Rise of Innovation Districts: A New Geography of Innovation in America", documented an emerging urban geography of innovation that sits at the intersection of economy-shaping, place-making, and network-building. **Innovation districts** are defined as "geographic areas where leading-edge anchor institutions and companies cluster and connect with start-ups, business incubators, and accelerators". They are also physically compact, transit-accessible, and technically wired and offer mixed-use housing, office, and retail. Today, more than 100 innovation districts are emerging around the world.

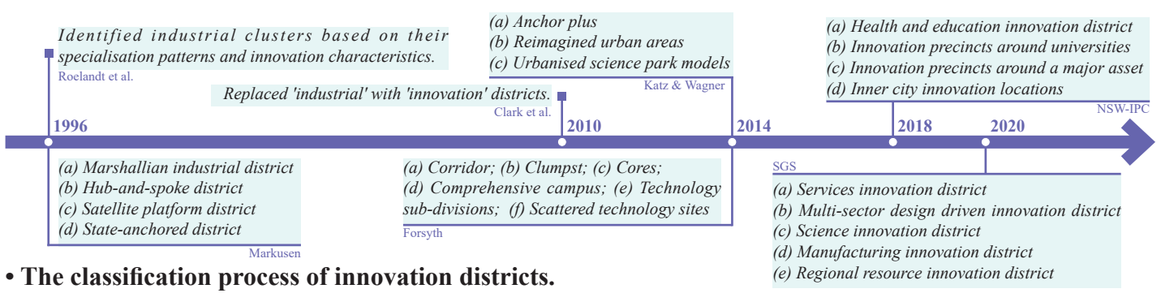
• Examples of globally recognised innovation districts.



1.2. Research gap and purpose

So far, several scholars have attempted to classify innovation districts based on features, functions, or spatial characteristics. Nonetheless, these classifications were mostly concluded from those in mature stages and often excluded spatial attributes. For China and other countries, where innovation districts are still an early trend, identifying potential types and conducting appropriate plans to incentivise the clustering of innovation activities are more urgent concerns.

Formalized innovation districts' design, implementation, and management is also a relatively new practice area. Much of the existing research provides a mix of insights and empirical justifications for investing in an innovation district, which could not instruct planners and policymakers to build an actual one. Therefore, there is a need to translate theories and descriptions of innovation districts into advice.



• The classification process of innovation districts.

The purpose of this project is to expand our understanding of the classification of potential innovation districts using clustering analysis and provide initial guidance to city leaders on how best to recognize and extend the growth of their own innovation districts, building on the distinctive assets.

2. Studied area and data

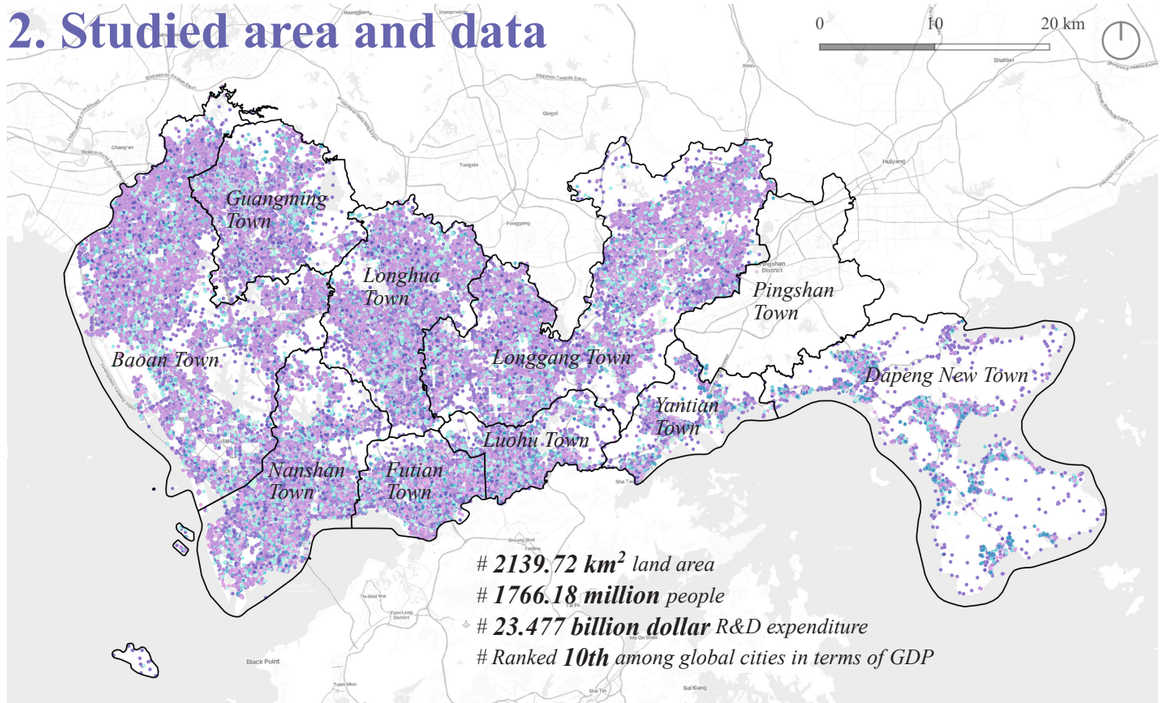


Fig. 1. Spatial distribution map with sample locations of POI data sets in Shenzhen.

2.1. Study area

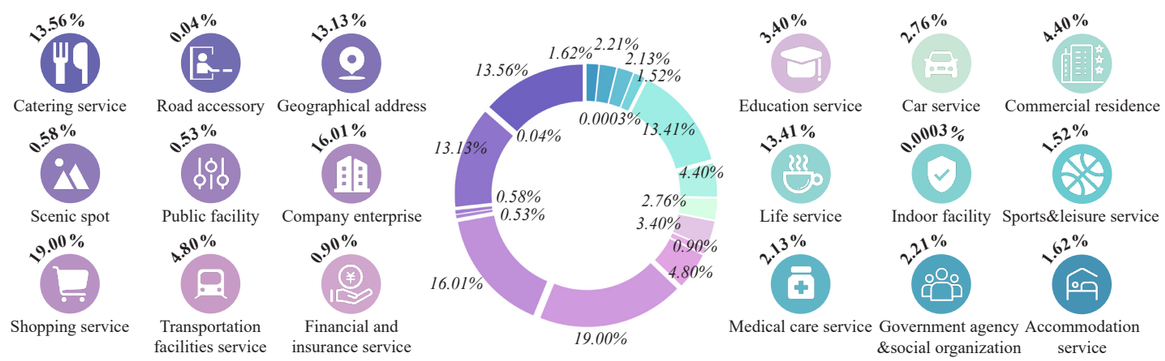
Over the last 40 years, Shenzhen has transformed itself into a global hub for technology development and secured the top position in the 2022 National Innovation City Innovation Capability Evaluation Report. The city's GDP increased from 0.196 billion Yuan in 1979 to 3238.768 billion Yuan in 2022 by over 16,000 times. Despite great economic achievement, the city now suffers from limited land resources and possible relocation of enterprises because of the rising living and production costs. Innovation districts, which are compact and decentralized, could be a perfect solution for city leaders of Shenzhen to drive growth in years to come.

2.2. Data preparation and software

As the city is a complex system containing numerous and miscellaneous elements, the following data needs to be prepared to identify potential innovation districts of Shenzhen: the basic grid map, and POI data of various kinds of enterprises and urban services.

The raw POI data are obtained from the Baidu Map website using Python crawlers, properties of which include name, latitude and longitude coordinates, functional category of the POI, etc. I use QGIS (ver. 3.32) to spatially visualize the various data, and use WGS_1984_UTM_Zone_49N as the unified coordinate system.

• POI data types and their volume distribution.



3. Methodology

3.1. Data pre-processing

Selection of variables

The clustering of innovation districts in the case city relies on establishing a distinct indicator system. A robust innovation ecosystem is contingent upon a solid foundation encompassing human capital, cultural assets, information technology infrastructure, and the interconnectedness of these elements. Considering the spatial attributes inherent to innovation districts, the physical factor system can be classified into three components: innovation source, innovation service, and innovation environment.

Table 1: Physical factor system of innovation districts

Innovation ecosystem	Innovation district	The determining factor	Description of a group of factors
Human capital	Innovation source	Anchor institute	Universities & Institutes Key laboratories Engineering technology research centers Engineering laboratories
		High-tech enterprise	Electronic information, optomechatronics, new materials, new energy, biological/medical technology, environmental protection, aerospace, marine engineering and nuclear applications firms
		Creative enterprise	Media and art, industrial design, architectural design, fashion creativity firms
Information technology infrastructure	Innovation service	Upstream service enterprise	Financial service enterprises Business service enterprises
		Downstream service enterprise	Manufacturing enterprises Logistics warehousing enterprises
Culture, values, and norms	Innovation environment	Cultural facility	Cultural exchange facilities Cultural experience facilities Cultural relics

Selection of POI data

The next task is to select the specific POI data representing the six variables of innovation districts.

Table 2: Flowchart of data filtering and integration

Delete data with keywords | Filter data with keywords

	Basetype	Subtype	Category	Filter
Anchor institute	2.21% 510/23048	Scientific research institution; School	Scientific research institution Institution of higher learning	Delete duplicates.
High-tech enterprise	37.10% 40274/108550	Company Well-known enterprise	Telecommunications company; Mechanical and electronic; Network technology; Pharmaceutical company; Well-known enterprise	"marketing", "warehouse", "operation", "service".
		Company Company enterprise	Telecommunications company Company Company enterprise	"science&technology", "network", "new materials", "biology", "medicine", "aerospace", "marine", "environmental protection", "photoelectric", "new energy", "electronics", "robotics", "automation", "circuit", "chip".(1)
				"factory", "consulting", "project".
Creative enterprise	6.99% 7592/108550	Company Company enterprise Well-known enterprise		Delete data with keywords in (1).
				"warehouse", "factory", "cold chain", "trailer", "container", "community", "construction", "data", "automobile", "trade", "business", "agriculture", "petroleum", "machinery", "equipment", "aquaculture", "building material".
				"design", "media", "creative", "art", "fashion", "culture", "clothing", "advertising", "gallery", "packaging", "jewelry", "landscape", "furniture", "printing".
		Company enterprise	Construction company; Advertising decoration	"creative industry park".
Upstream service enterprise	4.94% 5361/108550	Company; Company enterprise		"business", "commerce", "trade", "finance", "investment", "capital", "banking", "security", "insurance", "trust", "futures", "funds", "market management", "law", "legal affairs", "consulting", "exhibition".
				"psychology", "construction", "education", "health".
Downstream service enterprise	6.53%	Company	Commercial trade	
		Factory; Company	Factory; Metallurgy and chemical industry	
		Company; Company enterprise		"factory", "warehouse", "cold chain", "logistics", "container".
Cultural facility	1.17%	Tea house; Cafe	Tea house; Coffee shop; Pacific Coffee Company; Uejima coffee; Starbucks Coffee	
	2.38%	Museum; Science Museum; Art Gallery; Library; Exhibition Hall		"college", "university", "company", "children", "love".
	8.91% 918/10301	Movie theater	Theater; Concert Hall	"script kill", "cinema", "children's palace", "detective", farm".
		Entertainment venue	Pub	"hotel", "restaurant", "mansion", "winery", "OCT East".
	0.27%	Exclusive shop	Bookstore; Antique calligraphy and painting	"terminal", "picture book", "advertising", "printing", "stationery".
	7.65% 300/3921	Scenic spot Scenic spot related	Scenic spot; Red scenic spot; Memorial hall; Tourist attraction; Temples and Taoist temple	"ancestral hall", "temple", "archway", "turret", "ruins", "old site", "old residence", "ancient building", "ancient village".

3.2. Clustering analysis

Selection of clustering algorithms

Cluster analysis is the method of classifying samples into different groups or subsets, with all samples in the same group having relatively similar properties, and **K-means clustering** is a standard, non-hierarchical clustering with pre-specified clusters.

In this project, Euclidean distance was selected for the K-means algorithm. The function can be represented as follow:

$$J = \sum_{i=1}^{k} \sum_{j \in C_i} \left\| x_i - \mu_i \right\|^2$$

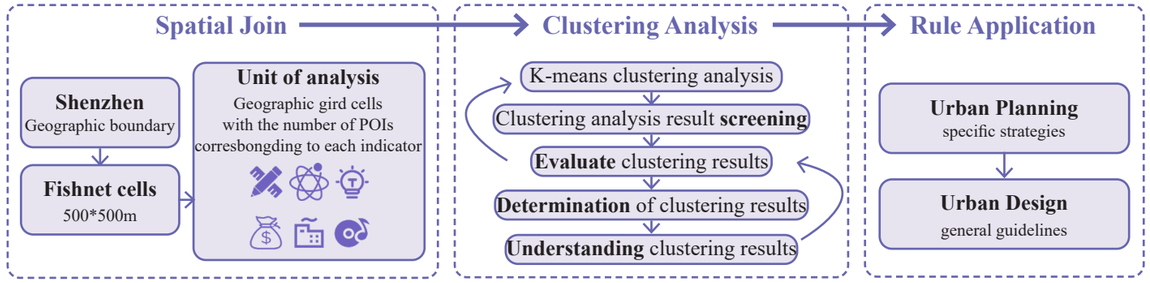
J the objective function
C_i the i^{th} cluster
n_i the number of samples in i^{th} cluster

$d_{ji} = ||x_j - \mu_i||^2$ the calculation of the distance between each sample point x_j and centroid μ_i in the i^{th} cluster

The centroid μ_i can be calculated based on the function below:

$$\mu_i = \frac{1}{|C_i|} \sum_{j \in C_i} x_j$$

The research process based on the K-means clustering algorithm can be summarized in the following steps:



Standardization of variables

The disparities in the initial magnitude of features may lead to an excessive influence of a feature set on the metric of the K-means algorithm. Consequently, **Feature Normalisation (FN)** is introduced to transform the distribution of elements along variables so that each has a mean of zero and a standard deviation of one.

The statistical FN methods analysed in this work are defined as:

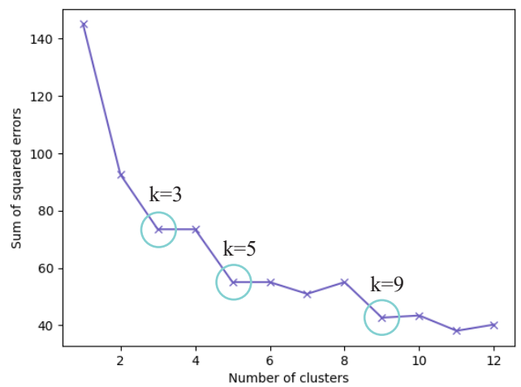
$$\widetilde{X}^{Norm} = \frac{X - \mathbf{pos(X)}}{\mathbf{dis(X)}}$$

pos(X) the position or central tendency statistic vector that centres the values of the features
dis(X) the dispersion statistic vector which scales the features

Determining the number of clusters

The **Elbow Method** was employed in this study to determine the optimal number of clusters. A marked flattening of the graph suggests that the combined clusters are highly dissimilar, thus the appropriate cluster number is found at the 'elbow' of the graph. According to Fig. 2, three distinct elbows are observed at positions three, five, and nine, indicating that multiple natural clusters fit the data well. Considering prior knowledge about data partitioning, it is plausible to assume that there are more than five types of innovation districts. Greater diversity in types necessitates more specific planning strategies, thereby enhancing future development prospects.

Fig. 2. The Elbow Method showing the optimal k.



Finally, this study selects the number of clusters as **nine**, sets the corresponding number of iterations, and performs the K-means clustering analysis using the 'Attribute Based Clustering' plugin (ver. 2.2.1) in QGIS.

4. Results and discussion

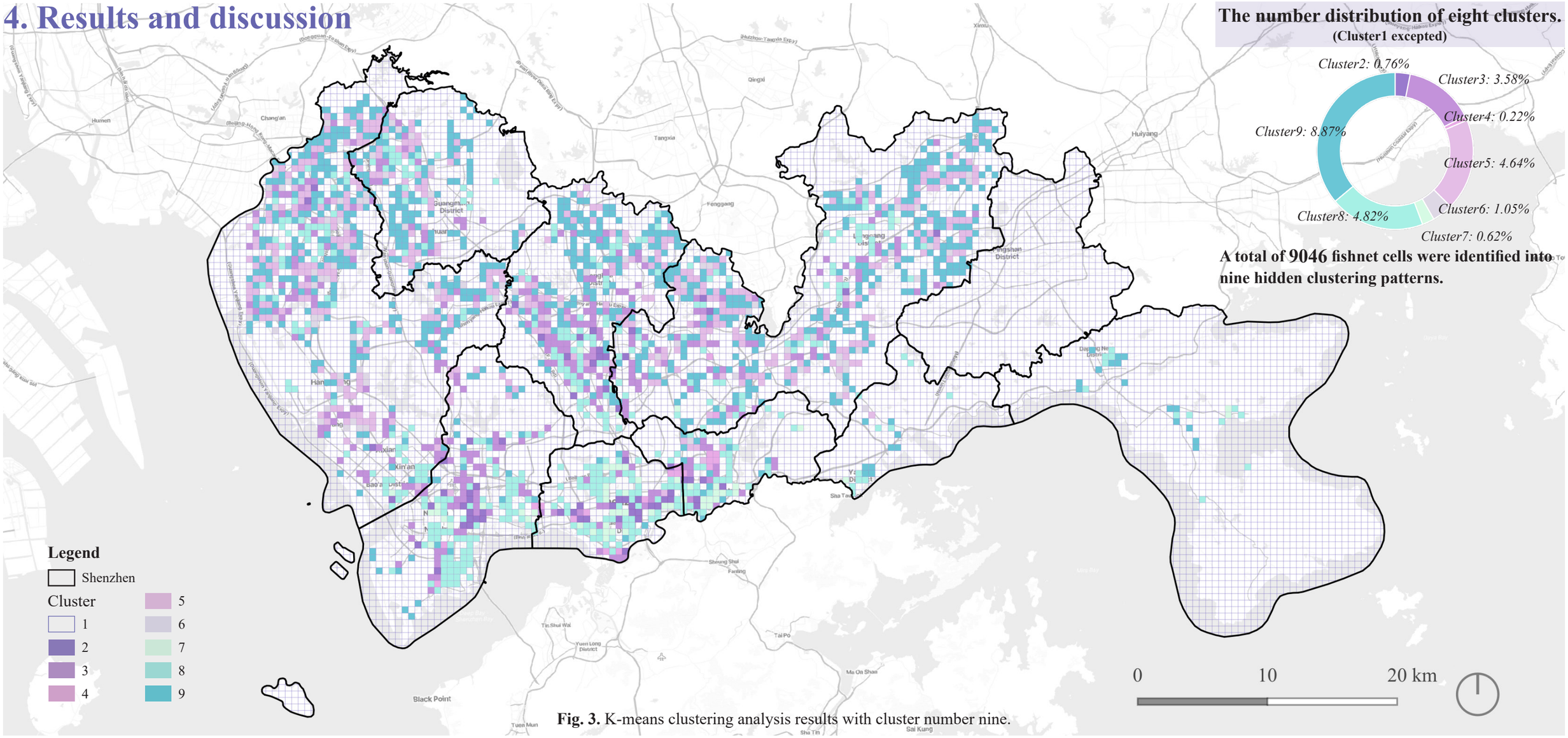
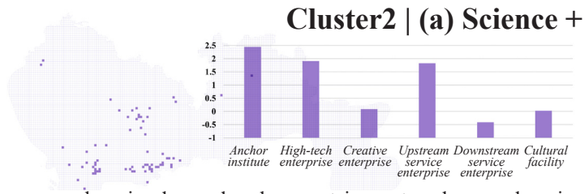


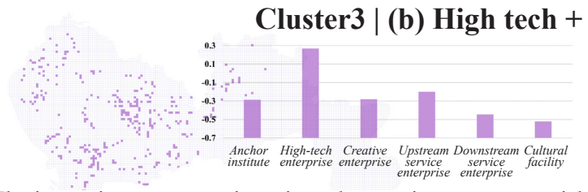
Fig. 3. K-means clustering analysis results with cluster number nine.

Cluster2 | (a) Science +



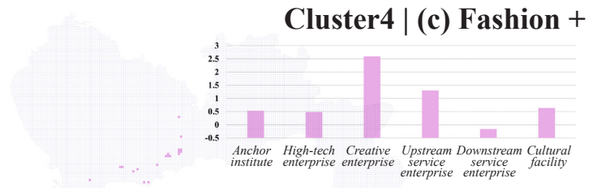
Large-scale mixed-use development is centered around major anchor institutions and a rich base of related firms, and spin-off companies involved in the commercialization of innovation.

Cluster3 | (b) High tech +



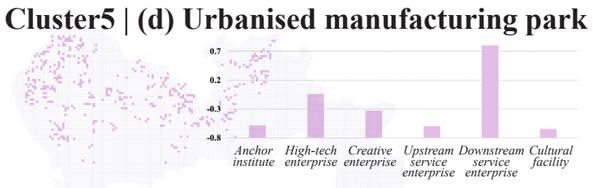
The inner city areas are witnessing a burgeoning entrepreneurial and innovation activity, primarily driven by the digital and financial technology sector and anchored by major high-tech companies.

Cluster4 | (c) Fashion +



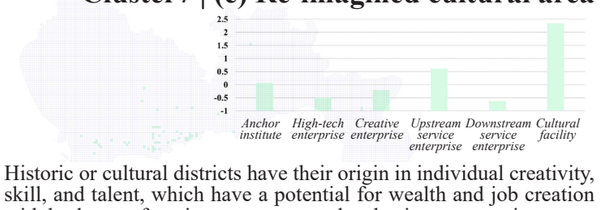
A clustering of complementary creative industries takes advantage of a concentration of industry talent and market demand to develop into a creative industries corridor and a fashion consumption place.

Cluster5 | (d) Urbanised manufacturing park



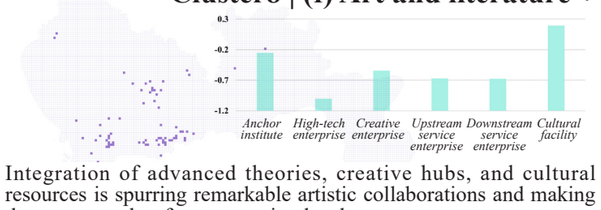
Industrial or warehouse districts are urbanizing through increased density and an infusion of new activities, which support collaboration and profit sharing to attract high-tech firms.

Cluster7 | (e) Re-imagined cultural area



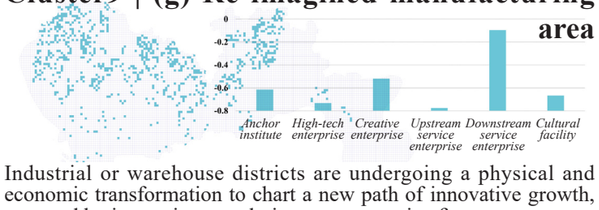
Historic or cultural districts have their origin in individual creativity, skill, and talent, which have a potential for wealth and job creation with brokerage functions to accommodate business expansion.

Cluster8 | (f) Art and literature +



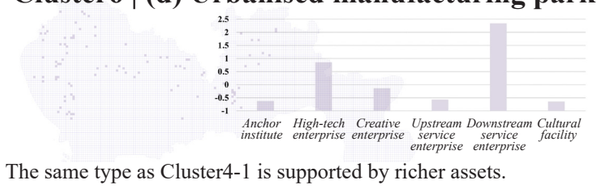
Integration of advanced theories, creative hubs, and cultural resources is spurring remarkable artistic collaborations and making the arts an anchor for community development.

Cluster9 | (g) Re-imagined manufacturing area



Industrial or warehouse districts are undergoing a physical and economic transformation to chart a new path of innovative growth, powered by increasing popularity among creative firms.

Cluster6 | (d) Urbanised manufacturing park

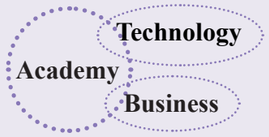
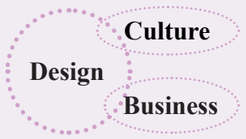
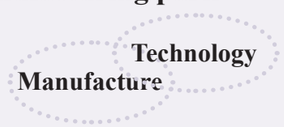


The same type as Cluster4-1 is supported by richer assets.

5. Asset-based strategies

As emphasized in the 2014 paper, thriving districts achieve their maximum potential by implementing asset-based strategies that capitalize on the **economic, physical, and social networking assets** within the district. This study has identified potential innovation districts by six essential characteristics using K-means cluster analysis, thereby highlighting the existing assets that should be taken into practical consideration. In Table 3 I formulate different strategy mixes which match each of the seven innovation districts identified previously.

Table 3: Asset-based strategies for seven types of potential innovation districts.

Types of Innovation Districts	Economic strategies Policymakers Investors	Spatial strategies Planners Designers	Networking strategies Organizers
(a) Science +  	• Sustained funding for scientific research • Investment in **incubators** • Provision of **affordable housing**	• Increase **cultural amenities** like music hall and coffee shops • Build public spaces between universities and companies • **Open amenities on campus** like playgrounds, gyms, and libraries to the public • Open ground floor lobbies of companies to the public	• Hold lecture series of researchers and entrepreneurs • Hold **startup competitions** to connect college students with investors • Offer job training for local unskilled labors
(b) High tech +  	• Relaxation of regulations to attract high-tech firms • Investment in **entrepreneurial support organizations** • Encourage **retails, hotels and restaurants**	• Build parks and green spaces for IT workers to relax • Build coworking spaces • Provide amenities like cafeterias, gyms, or lounges publicly	• Hold technological workshops • Hold free sports classes • Offer **science education** for local residents
(c) Fashion +  	• Sustained funding for **effective marketing** • Investment in **chambers of commerce** • Encourage **aesthetically** pleasing spaces	• Set tables, chairs, and benches where people can sit and walk • Build coworking spaces • **Open ground floor lobbies** of creative firms to the public for presenting products	• Hold art salons to gather designers and media • Hold creative markets for the public • Offer **art education** for local residents
(d) Urbanised manufacturing park  	• Relaxation of regulations to attract **high-tech firms** • Investment in incubators • Provision of affordable housing	• Build parks and green spaces to **ease pollution** • Revitalize factories to coworking places • Provide amenities like cafeterias, gyms, or lounges publicly	• Hold technological workshops • Hold free sports classes • Offer production training for local unskilled labors
(e) Re-imagined cultural area  	• Relaxation of regulations to attract **creative firms** • Investment in chambers of commerce • Encourage **fashionably** pleasing spaces	• Narrow streets with wider sidewalks • **Revitalize historical buildings** into public cultural places • Extend public transit to the district	• Hold design workshops • Hold **outdoor movie screenings** for the public • Offer **cultural education** for local residents
(f) Art and literature +  	• Relaxation of regulations to attract creative firms • Investment in incubators • Provision of **affordable housing**	• **Narrow streets with wider sidewalks** • Build coworking spaces • Open amenities on campus like playgrounds, gyms, and libraries to the public	• Hold art salons to gather students and artists • Hold creative markets for the public • Offer career-building workshops for low-income residents
(g) Re-imagined manufacturing area  	• Relaxation of regulations to attract **creative firms** • Investment in incubators • Encourage **retails, hotels and restaurants**	• Build parks and green spaces to improve the landscape • **Revitalize factories** into workshops or landmarks • Extend public transit to the district	• Hold design workshops • Hold **music festivals** for the public • Offer job training for local unskilled labors

6. Design guidelines

Design Principles

Innovation Districts

(a) Science +

(b) High tech +

(c) Fashion +

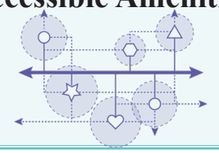
(d) Urbanised manufacturing park

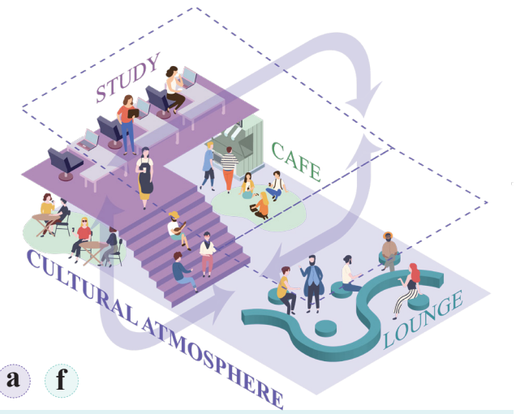
(e) Re-imagined cultural area

(f) Art and literature +

(g) Re-imagined manufacturing area

Provide Convenient and Accessible Amenities





a f

Integrate different **Experiences of Characters** related to the innovative and creative class

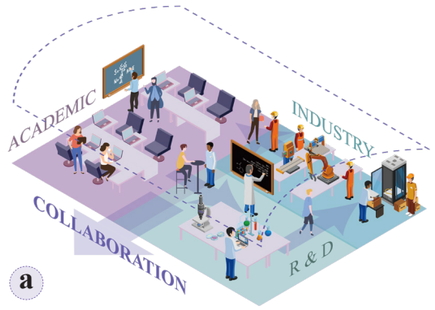
b d

Make a series of **Experiences of Uses** along the street and throughout the day and evening

e g

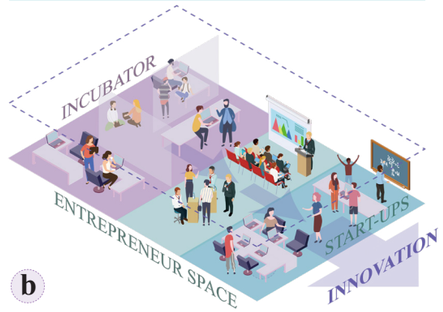
Be prepared to **Optimize Public Transport** options and consider changing current system

Build Coworking Places for Creative Collisions



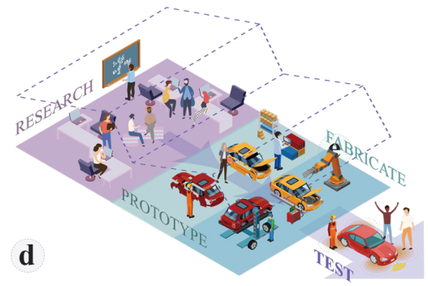
a

Multi-Disciplinary



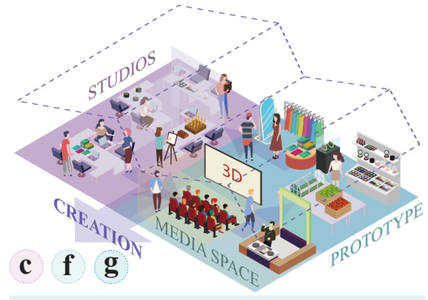
b

Ideation



d

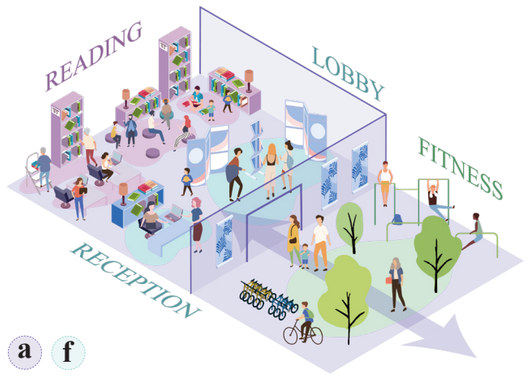
Research & Maker Hub



c f g

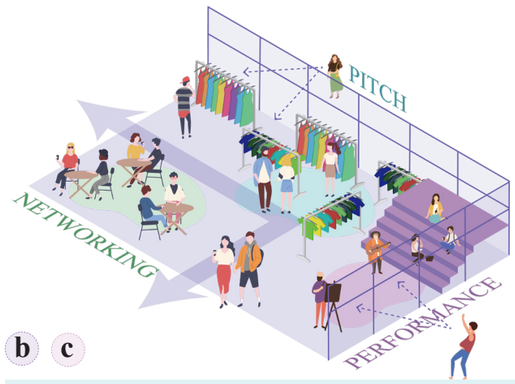
Design & Maker Hub

Share Lively Edges with Locals and the Public



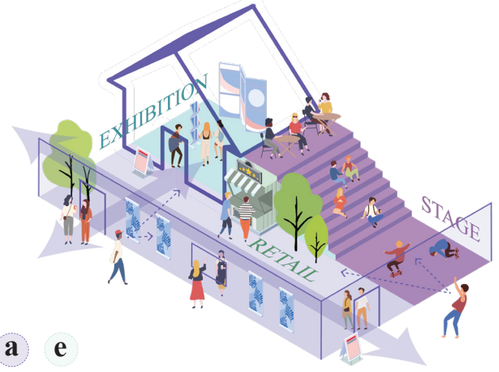
a f

Expand the functions of **Existing Public Spaces** to serve more Downtown Residents



b c

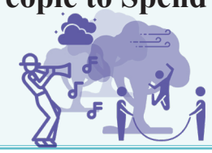
Open **Ground Floor Lobbies** of companies to stimulate commercial activities of the public



a e

Program adjacent **Underused Sites**, like historical buildings and private plazas

Create 24-Hour Invitations for People to Spend Time



b d

Improve the **Conditions for Staying** by creating a natural and lively public realm

c e f

Make a network of **Complete Streets** where it is possible to walk, bike, drive, and spend time

f g

Increase the **Variety of Activities** to attract a larger quantity and diversity of users