Case Studies on Transport Policy 13 (2023) 101055

Contents lists available at ScienceDirect

Case Studies on Transport Policy

journal homepage: www.elsevier.com/locate/cstp

Content analysis of Chinese cities' Five-Year Plan transport policy documents

Meng Guo [a,*], Qingyang Li [b], Chenyang Wu [c,d,*], Scott Le Vine [d,e], Gang Ren [a]

[a] School of Transportation, Southeast University, China
[b] Department of Industrial and Manufacturing Systems Engineering, University of Hong Kong, Hong Kong SAR, China
[c] School of Aeronautics, Northwestern Polytechnical University, China
[d] Urban System Lab, Imperial College London, United Kingdom
[e] Department of Geography, SUNY New Paltz, United States

ARTICLE INFO

Keywords:
Developing countries
Five-Year Plan
Content analysis
Sentiment analysis
China

ABSTRACT

Chinese cities' Five-Year Plans establish the direction and priorities for their urban transport systems for the subsequent half-decade. In this study, we compiled and then analyzed the document structure and content of the Five-Year Plans for Transport for 23 Chinese cities at two points in time (2010 and 2020). These documents afford a unique opportunity to establish how Chinese transport policies vary across different types of cities during a period of rapid economic growth and transport network development. Using content analysis techniques, we report findings regarding frequency and sentiment (positive/negative) towards a set of traditional and emerging transport planning topics, as well as document legibility (words per sentence, sentences per paragraph, use of graphical content, etc.). Our findings indicate that compared to many western Regional Transport Plans, Chinese Five-Year Plans tend to be more text-heavy than similar documents in western countries, with their length increasing during the 2010s. Furthermore, we demonstrate systematic variability in terms of topic frequencies with respect to cities' characteristics during the period 2010 to 2020, however the variations in terms of sentiment scores are weaker. More economically advanced cities exhibited a heightened focus on sustainable travel options, negative externalities related to cars, emerging technologies, and equity, whereas less economically developed cities emphasized regional development issues. We also demonstrate an increased focus over time on equity, vehicle technology, and parcel-delivery logistics within the Five-Year Plans. Of the various topics we analyzed, air pollution experienced the greatest decrease over time in sentiment.

1. Introduction

China has experienced rapid modernization in recent decades. In the transport sector, this has manifested as steep changes in indicators such as the quality and quantity of car infrastructure (Ecola et al., 2019), private car ownership (an increase from 44 per 1,000 population in 2010 to 172 per 1,000 in 2020) and length of the high-speed rail network (5,133 km in 2010 vs. 37,929 km in 2020) (National Bureau of Statistics of China [NBS], 2022). China also underwent rapid economic growth during this period (30,000 CNY in 2010 [~ 4,400 USD]; 72,000 CNY per capita in 2020 [~ 10,700 USD]), however economic development remains unevenly distributed within the country, with eastern cities generally characterized by higher levels of development (disposable income per capita in China's eastern regions of 41,000 CNY (6,100 USD) as of the year 2020, compared with 25,400 CNY (3,800 USD) in western regions (NBS, 2022)). As in many other developing societies, China's cities are subject to major transport-related challenges, including air quality, road safety, congestion, energy consumption, and regional accessibility.

The motivation for this line of research is to document themes within Chinese cities' Five-Year Plans for Transport (FYP-Ts) as well as how their content (textual and thematic) evolved during this period of rapid development (2010 – 2020). The data sample includes 23 cities encompassing all four of China's city "tiers" (first-tier, new first-tier, second-tier, and third-tier cities) and all of China's primary geographic regions.

* Corresponding authors at: School of Aeronautics, Northwestern Polytechnical University, China (C. Wu); School of Transportation, Southeast University, China (M. Guo).

E-mail addresses: guomeng@seu.edu.cn (M. Guo), cywu@nwpu.edu.cn, c.wu15@imperial.ac.uk (C. Wu).

https://doi.org/10.1016/j.cstp.2023.101055
Received 14 March 2023; Received in revised form 1 June 2023; Accepted 28 July 2023
Available online 29 July 2023

China's FYPs set high-level public policy direction for the development of different sectors, but do not directly allocate resources or control production (Fan, 2006; Heilmann, 2011). The most well-known FYP is the 'general' FYP, which focus on the high-level social and economic development objectives that cut across sectors (like 'becoming a moderate developed economy in 2035′ and 'achieving carbon neutrality by no later than 2060′). In contrast, FYP-Ts – the focus of this study – are separate documents that are limited in scope to the transport sector; analogous documents for different sectors are prepared by each other department of each Chinese municipal government (Education, Energy, Science/Technology, etc.). The first set of FYPs was established in 1953, and Plans have subsequently been updated every five years (Casey and Koleski, 2011). The current generation is the 14th FYPs, covering the period 2021 to 2025.

The process of preparing and adopting each of the various types of FYPs is broadly similar for FYP-Ts and other FYPs. All levels/branches of government (including the National Ministry of Transport and Bureaus/Commissions of Transport at the provincial and lower levels of government) complete FYP-Ts which summarize achievements over the past five years, establish target goals, identify problems/issues, and then set out specific tasks to be completed in the upcoming five years (Heikkila and Xu, 2014).

The workflow of the Five-Year Plans is nested downwards from the national level to lower levels of government. China's administration divisions have four levels, namely Provincial level (e.g., provinces and municipalities), Prefectural level (e.g., cities), County level (e.g., districts and counties), Township level (e.g., subdistricts and towns) and Village level (e.g., communities and villages). The FYP must consider the requirements of the administrative division as well as the National and higher-level administrative divisions' FYPs. After an initial consultation draft is prepared, it is circulated for advisory input from other commissions/ministries of the local authorities, lower-ranking administrative divisions, outside experts, and the public. The FYP is then finalized by the standing committee of the local authorities' 'People's Congress' and approved by the full 'People's Congress'.

The body of literature reviewing strategic planning documents internationally contains a growing number of studies focusing on metropolitan regions in developed nations. Though Chinese FYPs have been employed in studies of topics including carbon emission policies (Hu, 2016), energy policy (Yu et al., 2020; Yuan and Zuo, 2011), and work safety policy (Wang et al., 2018), FYP-T documents have attracted relatively little attention from the transport-policy research community. Despite the richness of the content of FYP-Ts and their importance in setting the policy and investment agenda for China's rapidly developing urban regions, there is a gap in the literature investigating their thematic and topical coverage.

To the best of the authors' knowledge, this study is a novel contribution that establishes patterns in the content and structure of Chinese cities' FYP-Ts, which document the evolution of the practice of transport planning in China. As noted above, the period 2010 – 2020 encompasses rapid development in urban transport systems in Chinese cities. The analytical framework we employed includes the use and customized training of a sentiment analysis approach, which allows identification of the sentiment (positive/negative) of the document authors' written references to each of a set of transport planning topics, in addition to the frequency at which the topics appear within the FYP-Ts and general document-level indicators.

The remainder of this paper is organized as follows: Section 2 discusses earlier literature relevant both methodologically and substantively, and Section 3 describes the process of selecting and compiling the database of FYP-Ts used in this study. Section 4 presents the content and sentiment analysis approaches, and Section 5 presents the results of the study. The paper concludes with a summary of the main findings and future research needs in Section 6.

2. Literature review

2.1. Content analysis

Content analysis is a systematic research method to analyze textual information in a standardized way that allow evaluators to make inferences about that information (U.S. General Accounting Office, 1996). The process generally includes four distinct stages (Bengtsson, 2016): *decontextualization* (during which units of 'meaning' are identified), *recontextualization* (during which portions of the text that are not relevant to the identified meanings are excluded), *categorization* (during which themes and categories are identified), and *compilation* (during which conclusions are drawn). Motivations for undertaking content analysis vary widely: researchers may seek to identify the intentions of the creator of the content, and/or to reveal patterns in the content.

Using data from text sources such as social media, periodicals, and policy documents, content analysis has been widely used to capture and interpret public perceptions on social issues (Ratanaburi et al., 2021; Singleton and Clifton, 2017; Zha et al., 2023) and analyze information communication and dissemination (Chan et al., 2023; Yavetz and Aharony, 2023). In the context of transport, content analysis has been applied to document and analyze public attitudes towards specific policies, identify factors influencing public decision-making, and

Table 1
Content analysis in transport studies.

Research	Topic	Region	Data used in content analysis
Casas and Delmelle (2017)	Public perceptions of bus rapid transit	Colombia (Cali)	Tweets (September 19–27, 2014)
Moody et al. (2019)	Mapping the motorization and urbanization characteristics and transport policy priorities of Chinese cities	China (287 Chinese cities for motorization and urbanization mapping; 44 cities for transport policy priorities identification)	CEIC database (motorization and urbanization features) Report on the Work of the Government (transport policy priorities identification)
Ratanaburi et al. (2021)	How stakeholder participation affects the quality of bicycle lanes	Thailand (Rattanakosin area)	Semi-structured interviews, online blogs, research and dissertations, laws and regulations, internal documentation, and photos
Li et al. (2021)	Themes, characteristic of policy tools, current status and future prospects in high-speed rail policy	China	National high-speed rail policy from 1990 to 2017
Wang et al. (2021)	Factors promoting local government in China to adopt car restriction policies	China (287 Chinese municipalities)	Car restriction policy documents
Lesteven and Thébert (2022)	Public perception of autonomous vehicles	France	News articles from 2012 to 2017 Tweets from Dec 1, 2017 to May 4 2018
Wu et al. (2022)	Identify whether young adults' attitudes towards the car have changed over time	U.S.	Top 40 songs in the Billboard Year-End Hot 100 singles from 1956 to 2015
Present paper	***Evolution of the practice of transport planning in China***	***China (23 cities)***	***12th and 14th FYP-Ts***

investigate priorities in policy-making. Table 1 summarizes relevant studies in this field, which draw on data from a range of sources including social media and press article texts, music lyrics, interviews, and formal public-sector planning documents.

The body of literature employing content-analysis techniques to evaluate public-sector strategic planning documents contains a growing number of studies focusing on developed-nation metropolitan regions, such as the US's Long-Range Transportation Plans (LRTPs) (Boisjoly and El-Geneidy, 2017; Lee and Sener, 2016; Manaugh et al., 2015; Singleton and Clifton, 2017), Canadian transportation plans (Bista et al., 2021; Boisjoly and El-Geneidy, 2017) and UK's local transport plans (Elvy, 2014). Chinese FYPs have been employed in studies of topics including carbon emission policies (Hu, 2016), energy policies (Yu et al., 2020; Yuan and Zuo, 2011), and work safety policies (Wang et al., 2018); Heikkila and Xu reviewed the content of Chinese cities' general FYPs (not FYP-Ts) to develop clusters of "prototypical" Chinese cities (Heikkila and Xu, 2014), as well as to evaluate municipal development policies (Xu and Heikkila, 2020). Though other studies have drawn on various other types of data sources to investigate specific transport-related topics in China (Li et al., 2021; Moody et al., 2019; Wang et al., 2021), FYP documents have attracted relatively little attention from the transport-policy research community. To the authors' best knowledge, Zhang and Bai (2017) is the sole study to employ FYPs to analyze a transport-related issue; this study explored the specific issue of alternative fuel vehicle adoption policies in three Chinese regions, however Zhang and Bai employed the cities' general FYPs rather than the FYP-Ts which form the basis of the present study (see discussion of these two types of documents in the Introduction).

2.2. Sentiment analysis

Sentiment analysis (SA) is a subfield of natural language processing that aims to extract subjective information from bodies of text. The objective of SA is typically to classify the sentiment (positive, negative, or neutral) of a piece of text, such as a social media posting (Da Silva et al., 2014; Kouloumpis et al., 2021), product review (Fang and Zhan, 2015; Yang et al., 2020), or news article (Rao et al., 2014; Souma et al., 2019).

SA techniques can be categorized into machine learning (ML) and lexicon-based approaches. ML techniques can be further classified into supervised and unsupervised methods, with the latter used when an initial set of documents suitable for use as training data is unavailable. Lexicon-based approaches rely on a sentiment lexicon, a collection of terms with a defined sentiment value for each term (Serrano-Guerrero et al., 2015). Various SA algorithms have been proposed (readers are referred to comprehensive literature reviews in Yadav and Vishwakarma (2020), Mäntylä et al. (2018), and Medhat et al. (2014)). In recent years, BERT (Bidirectional Encoder Representations from Transformers), a language model developed by Google AI (Devlin et al., 2018) has been applied for sentiment analysis (Jamil et al., 2022; Kumar and Malathi, 2022; Singh et al., 2021).

SA has been employed across different domains including marketing (Kauffmann et al., 2020; Pagolu et al., 2017), customer service (Kang and Park, 2014; Kumar et al., 2019), and political analysis (Kaya et al., 2012; Rill et al., 2014). In the field of transport, a wide range of topics have been the subject of published studies, including ride-hailing services (Morshed et al., 2021), public transport (Sari and Ruldeviyani, 2020), car use (Wu et al., 2022), and airline services (Gitto and Mancuso, 2017; Kwon et al., 2021; Martin-Domingo et al., 2019). During the COVID-19 pandemic, SA was used to assess the public opinions on the impact of COVID-19 on transport (Morshed et al., 2021; Park et al., 2022; Sari and Ruldeviyani, 2020). Sentiment analysis within transport studies have typically focused on public opinion (Martin-Domingo et al., 2019; Qi et al., 2020; Sari and Ruldeviyani, 2020) (Baj-Rogowska, 2017), however the sentiment of transport policymakers remains under-researched.

3. Data

The database of FYP-Ts consists of documents from 23 cities across each of China's four standard economic subregions (East, Central, West, and Northeast (NBS, 2011)) as well as each of the four standard 'tiers' of cities indicating the level of socio-economic development (Yicai Global, 2022). While some cities make their FYP-Ts publicly available from municipal websites, unlike in many Western countries, there are Chinese cities that do not make them readily available to the public. After compiling FYP-Ts for the years 2010 and 2020 from a combination of public websites and a process of direct inquiries[1] to municipal staff, the achieved sample is summarized in Table 2.

The 'city tier' classification system is widely used to categorize Chinese cities by level of economic development. Although not formally employed by the Chinese government, the "Ranking of cities' business attractiveness in China" published annually by Yicai (Yicai Global, 2022) is the standard city-tier classification used by researchers and practitioners (Chen and Woo, 2020; Jia et al., 2020; Ma et al., 2021; Rogoff and Yang, 2022; Wen et al., 2022). The latest-available city tiers (for year 2022) define 4 'first-tier' cities, 15 'new first-tier' cities, 30 'second-tier' cities, 70 'third-tier' cities, 90 'fourth-tier' cities and 128 'fifth-tier' cities. The list of first-tier cities (Beijing, Shenzhen, Shanghai, Guangzhou) generally remains stable, however, the composition of cities in other tiers varies more over time.

4. Methods

This section summarizes the approach to analyzing the structure and content of the 2010 and 2020 FYP-Ts. The structure of the FYP-T documents was characterized by the following indicators:

- Total count of Chinese-language characters
- Number of pages
- Characters per sentence
- Sentences per paragraph
- Paragraphs per page
- Figures per 100 pages
- Tables per 100 pages

A total of 18 transport planning topics were identified as relevant to this study. In order to characterize the content of the FYP-Ts, two members of the research team independently reviewed each of the FYP-T documents to identify the relevant references (referred to interchangeably hereafter as 'tokens').

For discussion purposes, the set of tokens was grouped into three categories: 'travel options', 'challenges', and 'emerging trends'. Following a reconciliation process among the two members of the study team that had independently identified tokens, concurrence was reached on the specific words/phrases that would trigger the identification of a token (see Table 3). A total of 6,557 references were identified across the 46 FYP-T documents.

We employ the language model *Chinese-BERT-wwm (wwm acronym: whole word masking)*, which is derived from the original BERT model (Cui et al., 2021) developed, to identify the sentiment (positive vs. negative) of the references to the various transport planning topics. In the present study, we customized[2] the pre-trained model with a training

[1] The following FYP-Ts were obtained from direct inquiry: Guangzhou (2010), Wuhan (2010), Changsha (2010), Harbin (2010), Baoji (2010 and 2020), Mianyang (2010 and 2020), and Hanzhong (2010 and 2020). All other FYP-Ts were obtained from municipal websites. The authors have archived all FYP-T documents obtained for use in this study, which are now available for use by the research community at https://doi.org/10.17605/OSF.IO/2FZV3.

[2] The authors have archived the code for sentiment analysis at https://github.com/ChenyangWu1991/Five-Year-Plan.git

Table 2
Characteristics of cities included within data sample.

City	Region	2020 area (km^2)	2020 Gross Regional Product (GRP) per capita (CNY)	2020 population (1,000 people)	2019 Car ownership per 1,000 people	2019 Rail passenger trips per year (1,000 people)
First-tier cities						
Beijing	East	16,410	164,889 (+63%)	21,893 (+12%)	227 (+19%)	147,550 (+66%)
Shenzhen	East	1,997	159,309 (+12%)	17,634 (+70%)	195 (+44%)	87,410 (+275%)
Shanghai	East	6,341	155,800 (+54%)	24,884 (+8%)	129 (+186%)	128,340(+111%)
Guangzhou	East	7,434	135,047 (-2%)	18,740 (+48%)	106 (+18%)	145,300(+55%)
New first-tier cities						
Hangzhou	East	16,853	136,617 (+47%)	11,965 (+38%)	200 (+84%)	88,740 (+259%)
Wuhan	Central	8,569	131,441 (+67%)	12,326 (+26%)	236 (+344%)	188,077 (+158%)
Changsha	Central	11,819	123,297 (+39%)	10,048 (+43%)	226 (+236%)	60,490 (+268%)
Hefei	Central	11,445	108,427 (+48%)	9,370 (+26%)	195 (+528%)	46,300 (+192%)
Tianjin	East	11,967	101,614 (+4%)	13,866 (+7%)	174 (+99%)	53,320 (+101%)
Chengdu	West	14,335	85,679 (+32%)	20,938 (+49%)	209 (+111%)	1,489,967 (+990%)
Second-tier cities						
Dalian	Northeast	13,739	94,685 (-9%)	7,451 (+11%)	185 (+205%)	23,620 (+10%)
Guiyang	West	8,043	72,246 (108%)	5,987 (+39%)	257 (+117%)	12,640 (+20%)
Wenzhou	East	12,110	71,766 (+44%)	9,587 (+5%)	226 (+205%)	28,394 (+376%)
Nanning	West	22,245	54,669 (+60%)	8,742 (+31%)	169 (+318%)	37,320 (+270%)
Harbin	Northeast	53,076	54,570 (+11%)	10,001 (-6%)	163 (+355%)	48,020 (+25%)
Third-tier cities						
Wuhu	Central	6,004	102,964 (+60%)	3,644 (+3%)	144 (+524%)	11,672 (+167%)
Yancheng	East	16,931	88,731 (+110%)	6,710 (-8%)	142 (+420%)	1,970 (+8%)
Putian	East	4,143	82,753 (+106%)	3,211 (+15%)	108 (N/A)	8,400 (+733%)
Zibo	East	5,965	78,092 (-8%)	4,704 (+4%)	205 (+149%)	N/A
Baoji	West	18,117	67,666 (+94%)	3,322 (-11%)	100 (+499%)	N/A
Mianyang	West	20,248	55,049 (+106%)	4,868 (+6%)	134 (+248%)	10,788 (+32%)
Hanzhong	West	27,096	49,179 (+151%)	3,211 (-6%)	88 (+408%)	N/A
Fuyang	Central	10,118	34,399 (+185%)	8,200 (+8%)	83 (+619%)	12,937 (+52%)

Note 1: 1 Chinese Yuan (CNY) to USD exchange rate is ~$0.15 at the time of writing.
Note 2: The growth rate of GRP per capita has taken inflation into account.
Note 3: Values in parentheses are the growth rate relative to 2010. For example, 21,893 (+12%) means the population of Beijing in 2020 is 21,893 k (21.9 million) people, which is a + 12% increase from 2010.
Note 4: The links to the GRP per capita, population, car ownership, and rail passenger trips per year data have been archived at https://doi.org/10.17605/OSF.IO/2FZV3.
Note 5: To avoid the impacts of COVID, we use car ownership and rail passenger trips data from 2019.
Note 6: Putian did not open the 2010 car ownership data; Zibo, Baoji and Hanzhong did not open the 2019 rail passenger trip data.

set of 600 randomly selected references (a ~ 3% sample) which had been independently manually scored (-1 = negative, 0 = neutral, +1 = positive) by two human analysts of the research team. This trained BERT algorithm was then used to quantify the sentiment of the full set of tokens from the FYP-Ts.

5. Results

In this section, we first present the analysis of the FYP-T document structure, followed by token-frequency analysis and analysis of the sentiment of the various tokens. Due to the effort per city to acquire documents and evaluate the content of each FYP-T, the achieved sample encompasses 46 FYP-T documents (one each for years 2010 and 2020 for each of the 23 cities). Results from this small sample should therefore be viewed as indicative, and future research to corroborate the findings from this study, possibly using automated classification/identification methods, would be valuable.

5.1. Document structure

Table 4 summarizes the FYP-Ts' structures, showing that overall the total character counts of the FYP-Ts increased by 44% from 2010 to 2020. The number of pages also increased (27%) but not as rapidly as total character counts. In general, the FYP-Ts tend to be text-heavy documents with little graphical content (13 figures per 100 pages in 2010 and 1.4 figures per 100 pages in 2020) in comparison to many contemporary Regional Transport Plans in western societies (see e.g. New York Metropolitan Transportation Council (2021), Association of Bay Area Governments and Metropolitan Transportation Commission (2021), and Miami-Dade Transportation Planning Organization (2019)).

The number of Chinese-language characters per sentence of FYP-Ts averages approximately 60, which for comparison purposes is approximately double the average sentence length in mass-media Chinese newspapers (~30 characters per sentence, see Huang and Wang (2019)). Also, the number of sentences per paragraph increased by 54% between 2010 and 2020. Taken together, these metrics suggest difficulty for members of the general public to easily comprehend the content of Chinese FYP-Ts: FYP-Ts are text-heavy, are substantially more complex documents to read than newspaper articles, and became increasingly complex between 2010 and 2020. Various approaches to render FYP-Ts more accessible to general readership could include shorter/simpler explanatory articles (The State Council of the People's Republic of China, 2021a), interactive online programs (The State Council of the People's Republic of China, 2021b) and accompanying video-format content (Xinhua Net, 2021).

5.2. Token frequency

Due to the increasing salience of problems caused by automobile use in Chinese cities, we sought to investigate whether there has been an increase over time in the frequency of tokens associated with sustainable travel options ('bike', 'walking' and 'shared mobility'), as well as

Table 3
Listing of tokens and triggering words/phrases.

Token	Words/phrases that trigger reference identification
Travel options	
Bike	Bike; Riding; Active travel; Non-motorized vehicles; Non-motorized lanes; Greenway
Car	Car; Motor vehicle ownership; Private motor vehicle
Multimodality	Multimodality; Seamless/Efficient transfer in a journey; Integrated passenger stations that include more than one mode of transport; Integration of transport facilities
Shared mobility	Shared mobility; Bikesharing; Ridehailing; Customized public transport
Park & ride	Park & ride; P&R; P + R
Public transport	Transit; Public transport; Bus; Metro; Rail; Passenger station; Passenger motor vehicle
Walking	Walking; Pedestrian; Active travel; Non-motorized paths; Greenway
Challenges	
Air pollution	Air pollution; Emission; Emission standard
Climate change	Climate change; Low-carbon; Greenhouse gas emissions
Congestion	Congestion; Contradiction between supply and demand on urban roads
Equity	Equity for disadvantaged groups; Accessibility of public facilities to disabled people
Parking	Parking; Parking lots; On-street parking zone; Static traffic
Regional development balance	Equity among different regions; Reduce development gap between urban and rural areas; Promoting the development of countryside or surrounding areas
Safety	Traffic safety; Accident; Crash; Fatality and injuries
Emerging trends	
Alternative fuel vehicles	Alternative fuel vehicles; Clean energy vehicles; Hydrogen fuel
Autonomous vehicles	Autonomous vehicles; Autonomous techniques for vehicles
Electric vehicles	Electric vehicles; Charging stations
Parcel delivery	Parcel; Parcel delivery station; E-commerce logistics

concerns related to parking, air pollution, climate change, and traffic congestion. Furthermore, with the development of emerging mobility services and the energy industry, we sought to identify whether there has been a heightened focus on tokens belonging to the 'emerging trends' category ('parcel delivery', 'alternative fuel vehicle', 'autonomous vehicles', etc.).

Table 5 presents the frequency of tokens by city tier and year. For tokens belonging to the 'travel options' category, there is an increase in the frequency of 'bike', 'walking' and 'shared mobility'. 'Public transport' appears with the highest frequency in both years and across all city tiers. However, its frequency decreased in the top three tiers of cities (first-tier, new first-tier, and second-tier) while increasing in the lowest tier of cities (third-tier). This may be due to the inclusion of major railway construction plans in the 14th Five-Year Plans of many Chinese third-tier cities, in contrast to the relatively more mature transport networks of the higher-GRP Chinese cities.

For tokens categorized under 'challenges', the most frequently mentioned is 'regional development balance'. As the level of development of first-tier cities (all in eastern China, near the Pacific coast) is much higher than the rest of the country, the government has placed emphasis on promoting the development of parts of the country (both rural hinterlands of highly-developed cities and less-developed cities in other regions) through 'balanced development'. This finding is therefore reasonable, and notably there has been a decreasing trend in the frequency of 'regional development balance' in all tiers' FYP-Ts, which is found to be especially strong in the top three tiers of cities.

The frequency of 'climate change', 'air pollution' and 'parking' increased, particularly in the first-tier cities, which are in line with our expectations. The frequency of congestion, however, decreased in the first- and new-first tier cities. References to 'equity' for disadvantaged populations increased in all tiers except the second, and most sharply for first-tier cities (see the frequency changes of 'equity' in Table 5), suggesting that transport planners (particularly in China's most economically advanced cities) are increasingly focusing on addressing the needs of the disadvantaged population groups.

Finally, as we had anticipated, the frequency of all tokens in the 'emerging trends' category increased. 'Parcel delivery' experienced the most rapid increase, which is in accordance with the rapid growth of China's parcel delivery industry during this timeframe (Liang and Tan, 2019; NBS, 2022). The increased frequency of 'electric vehicles' is smaller than references to 'alternative fuel vehicles', suggesting a broadening in the discussions regarding alternative propulsion technologies. 'Autonomous vehicles' is the only token that did not appear in the 2010 FYP-Ts. In 2020, reference to 'autonomous vehicles' is much more frequent in the first-tier and the new first-tier cities rather than in the two lower city tiers.

5.3. Token frequency and city characteristics correlations

In this section, we evaluate how city characteristics influence the frequency of different topics through the correlation between city characteristic indicators mentioned in Table 2 and token frequency. As the externalities caused by cars are likely to be more serious in larger and more economically developed cities, we wish to explore whether the frequency of tokens related to sustainable forms of travel (e.g. bike and walking), challenges arising from car use (e.g. climate change and congestion), and emerging trends (e.g. electric vehicles) will exhibit positive correlations with the various city-level indicators (population, GRP per capita, car ownership, and rail passenger journeys).

Table 6 presents the correlations between the geodemographic (area, population and GRP per capita) and transport indicators (car ownership and rail passenger trips). With three exceptions (car ownership and GRP per capita in 2010; population and rail passenger trips in 2010; and population and GRP per capita in 2020), the absolute values of the correlations are below 0.5. As a result, we proceeded to correlate each of the indicators in Table 6 with the token frequencies. However, the correlations between area (km^2) and token frequencies are very weak, with only the frequency of 'autonomous vehicles' showing a significant correlation with area (-0.39, $p=0.06$) at the 90% level of significance. Therefore, the correlation analysis results for area are excluded from this point forward.

The correlations between city characteristics and token frequencies are presented in Table 7. Most tokens correlate positively with the four indicators, but the correlations tend to strengthen from 2010 to 2020 for the two geodemographic indicators and weaken for the two transport indicators. Sustainable travel options ('bike', 'walking', and 'public transport'), challenges ('congestion' and 'parking') and emerging trends ('electric vehicles') are referenced more frequently in cities with higher population and GRP/capita in both years; the correlation between 'equity' and the two geodemographic indicators is only significant in 2020. However, except for 'congestion' which maintains a strong correlation with all indicators across most years, the correlation between other tokens related to sustainable travel, challenges, and emerging trends are generally weaker with the transport indicators than with the geodemographic indicators.

Given the detailed discussion on regional development balance issues in the FYP-Ts of less developed cities, 'regional development balance' is the only token that correlates negatively with the indicators. Similar to most other tokens, its correlation is much stronger for the two geodemographic indicators. These findings align the observations from Table 5, where references to 'regional development balance' can be seen to be higher in the lower-tier cities.

5.4. Sentiment analysis

With respect to the sentiment analysis results, we wished to evaluate

Table 4
Statistical summary of FYP-T documents' structure.

	Total character count		Number of pages		Characters per sentence		Sentences per paragraph		Paragraphs per page		Figures per 100 pages		Tables per 100 pages	
	2010	**2020**	**2010**	**2020**	**2010**	**2020**	**2010**	**2020**	**2010**	**2020**	**2010**	**2020**	**2010**	**2020**
First-tier cities														
Beijing	16,431	34,768	36	81	85	45	1.35	3.31	3.97	2.85	0.00	0.00	1.18	26.67
Shenzhen	16,323	32,364	52	75	47	44	2.84	3.91	2.35	2.49	71.15	0.00	14.89	20.36
Shanghai	19,063	18,336	*	38	44	46	2.71	2.81	*	3.76	*	0.00	*	2.17
Guangzhou	10,213	29,235	23	65	53	58	2.24	2.74	3.74	2.83	0.00	0.00	3.77	18.97
New first-tier cities														
Hangzhou	21,650	37,770	45	65	70	73	1.80	4.39	3.82	1.80	0.00	0.00	10.00	13.61
Wuhan	18,401	16,440	35	30	61	59	1.74	3.54	4.94	2.60	0.00	0.00	1.64	3.36
Changsha	60,148	38,076	119	70	58	63	2.06	3.05	4.23	2.83	22.69	0.00	118.73	6.33
Hefei	32,512	33,728	89	78	57	51	1.98	4.09	3.24	2.06	15.73	5.13	8.77	45.10
Tianjin	10,384	27,090	47	47	50	54	1.60	3.22	2.79	3.32	85.11	0.00	14.00	0.00
Chengdu	11,375	37,320	23	102	47	70	2.92	3.18	3.61	1.64	0.00	3.92	0.00	25.71
Second-tier cities														
Dalian	12,038	21,812	20	41	75	50	2.04	5.27	3.95	2.02	0.00	0.00	9.33	0.00
Guiyang	14,351	30,194	*	*	67	65	1.73	2.12	*	*	*	*	*	*
Wenzhou	30,941	30,296	62	52	71	58	1.90	4.02	3.71	2.50	0.00	0.00	22.54	19.23
Nanning	19,905	27,195	40	44	56	58	2.50	2.96	3.55	3.55	0.00	0.00	5.35	3.40
Harbin	20,213	71,786	60	142	70	73	1.42	2.01	3.38	3.41	33.33	14.79	31.46	40.58
Third-tier cities														
Wuhu	26,359	28,702	36	68	62	50	1.64	2.07	7.17	4.01	0.00	0.00	9.61	15.75
Yancheng	14,960	23,938	*	49	52	52	2.18	3.66	*	2.55	*	0.00	*	7.64
Putian	42,537	25,977	99	57	77	67	2.35	3.15	2.35	2.14	5.05	0.00	34.78	8.87
Zibo	19,478	11,391	46	23	55	46	2.27	3.36	3.33	3.17	15.22	0.00	17.87	0.00
Baoji	14,325	19,370	37	45	59	58	1.67	3.39	3.95	2.20	0.00	0.00	15.25	24.14
Mianyang	18,404	22,858	31	52	61	49	2.12	2.72	4.61	3.33	0.00	1.92	0.00	14.29
Hanzhong	15,130	21,687	31	42	52	62	2.53	1.58	3.71	5.26	0.00	2.38	0.00	19.35
Fuyang	13,000	47,854	*	104	68	63	1.81	2.68	*	2.69	*	2.88	*	12.54
Average	**20,789**	**29,921**	**49**	**62**	**61**	**57**	**2.06**	**3.18**	**3.81**	**2.86**	**13.07**	**1.41**	**16.80**	**14.91**

Note: * identifies FYP documents published as single-scrolling web pages (i.e., no page numbering).

Table 5
Frequency of tokens for different tiers (per 1,000 characters).

	First tier		New first tier		Second tier		Third tier		Average	
	2010	2020	2010	2020	2010	2020	2010	2020	2010	2020
Travel options										
Public transport	5.16	4.26	4.60	3.20	3.78	3.22	1.86	2.83	3.57	3.26
Multimodality	1.29	1.03	0.81	0.68	1.07	0.64	0.19	0.44	0.73	0.65
Shared mobility	0.32	0.34	0.08	0.38	0.01	0.30	0.00	0.15	0.08	0.28
Bike	0.37	0.74	0.28	0.20	0.06	0.13	0.00	0.09	0.15	0.24
Walking	0.29	0.65	0.29	0.20	0.08	0.14	0.01	0.07	0.15	0.22
Car	0.32	0.22	0.23	0.14	0.09	0.06	0.05	0.02	0.15	0.09
Park & ride	0.10	0.09	0.06	0.02	0.02	0.01	0.00	0.01	0.04	0.03
Challenges										
Regional development balance	0.45	0.36	0.81	0.69	1.21	0.83	1.18	1.06	0.96	0.79
Parking	0.62	0.86	0.36	0.12	0.17	0.24	0.01	0.07	0.24	0.26
Climate change	0.28	0.55	0.31	0.25	0.14	0.09	0.06	0.22	0.18	0.26
Air pollution	0.15	0.24	0.26	0.21	0.17	0.17	0.08	0.16	0.16	0.19
Safety	0.28	0.20	0.12	0.17	0.19	0.15	0.13	0.15	0.17	0.17
Congestion	0.52	0.31	0.37	0.18	0.13	0.14	0.04	0.04	0.23	0.14
Equity	0.05	0.29	0.00	0.05	0.02	0.02	0.01	0.05	0.02	0.09
Emerging trends										
Parcel delivery	0.15	0.54	0.04	0.91	0.12	0.90	0.14	0.94	0.11	0.85
Alternative fuel vehicle	0.19	0.49	0.06	0.25	0.10	0.26	0.06	0.30	0.09	0.31
Electric vehicles	0.18	0.35	0.07	0.13	0.06	0.19	0.01	0.09	0.07	0.17
Autonomous vehicles	0.00	0.16	0.00	0.16	0.00	0.09	0.00	0.03	0.00	0.10

whether sentiment towards sustainable travel options (bike, walking, shared mobility, etc.) and electric vehicles became more positive between 2010 and 2020, and whether sentiment towards cars and the negative externalities associated with cars (parking, air pollution, congestion, etc.) have become more negative. With the objective of generating a second set of sentiment analysis results that are independent of the BERT algorithm for confirmation purposes, two of the research team's human analysts also (independently of both the BERT

Table 6
Correlation between geodemographic and transport characteristics of cities.

	2010					2020				
	Area (km^2)	**Population (1,000 people)**	**GRP per capita (CNY)**	**Car ownership per 1,000 people**	**Rail passenger trips per year (1,000 people)**	**Area (km^2)**	**Population (1,000 people)**	**GRP per capita (CNY)**	**Car ownership per 1,000 people**	**Rail passenger trips per year (1,000 people)**
Area (km^2)	1.0	0.209 (0.34)	−0.376 (0.08)	−0.236 (0.28)	−0.013 (0.96)	1.0	−0.149 (0.50)	−0.437 (0.04)	−0.133 (0.54)	−0.024 (0.91)
Population		1.0	0.159 (0.47)	0.210 (0.34)	**0.637 (<0.01)**		1.0	**0.703 (<0.001)**	0.248 (0.25)	**0.485 (0.02)**
GRP per capita			1.0	**0.601 (<0.01)**	**0.449 (0.03)**			1.0	**0.358 (0.09)**	0.078 (0.73)
				1.0	**0.468 (0.02)**				1.0	0.218 (0.32)
					1.0					1.0

Note 1: In the format X (Y), X denotes correlation, and Y denotes p-value.
Note 2: We used 2019 car ownership and rail passenger trip data to avoid the impact of COVID on the two indicators.

Table 7
Correlation between frequency of tokens and characteristics of the cities.

	Population (1,000 people)		GRP per capita (CNY)		Car ownership per 1,000 people		Rail passenger trips per year (1,000 people)	
	2010	**2020**	**2010**	**2020**	**2010**	**2020**	**2010**	**2020**
Travel options								
Bike	*0.47 (0.02)*	**0.60 (<0.01)**	*0.45 (0.03)*	**0.70 (<0.01)**	*0.43 (0.04)*		*0.49 (0.02)*	
Car	*0.43 (0.04)*	**0.62 (<0.01)**	0.41 (0.05)	**0.57 (<0.01)**	*0.41 (0.05)*	0.40 (0.06)		
Multimodality	0.41 (0.05)	**0.54 (<0.01)**	**0.59 (<0.01)**	**0.65 (<0.01)**	*0.48 (0.02)*			
Shared mobility	**0.66 (<0.01)**	0.39 (0.06)			0.38 (0.08)		**0.53 (<0.01)**	
Park & ride	**0.67 (<0.01)**	*0.44 (0.03)*	*0.42 (0.05)*	*0.51 (0.01)*	*0.48 (0.02)*		*0.47 (0.02)*	
Public transport	0.39 (0.06)	**0.57 (<0.01)**	**0.64 (<0.01)**	**0.58 (<0.01)**	*0.46 (0.03)*	0.37 (0.09)	**0.54 (<0.01)**	
Walking	0.37 (0.08)	**0.59 (<0.01)**	0.41 (0.05)	**0.70 (<0.01)**	*0.42 (0.05)*		*0.49 (0.02)*	
Challenges								
Air pollution		0.35 (0.10)						
Climate change		**0.54 (<0.01)**	*0.48 (0.02)*	0.38 (0.08)	*0.50 (0.02)*		*0.43 (0.04)*	
Congestion	**0.59 (<0.01)**	**0.57 (<0.01)**	*0.52 (0.01)*	**0.62 (<0.01)**	**0.73 (<0.01)**	**0.54 (<0.01)**	**0.73 (<0.01)**	
Equity		**0.66 (<0.01)**		**0.59 (<0.01)**				
Parking	**0.69 (<0.01)**	**0.62 (<0.01)**	0.38 (0.08)	**0.56 (<0.01)**	**0.59 (<0.01)**		**0.56 (<0.01)**	
Regional development balance		**−0.67 (<0.01)**	**−0.55 (<0.01)**	*−0.43 (0.04)*	−0.40 (0.06)			
Safety					*0.48 (0.02)*			
Emerging trends								
Alternative fuel vehicle		*0.46 (0.03)*	0.39 (0.07)	0.36 (0.09)	**0.55 (<0.01)**			
Electric vehicles	*0.47 (0.02)*	**0.57 (<0.01)**	0.37 (0.08)	*0.47 (0.03)*				
Autonomous vehicles	–		–	*0.42 (0.05)*	–		–	
Parcel delivery								

Note 1: Correlations with *p*-value > 0.10 were excluded from the table.
Note 2: There are no references to 'autonomous vehicles' in the 2010 FYP-Ts.
Note 3: Bold and italics indicate significance at the 0.01 and 0.05 levels, respectively.
Note 4: We used 2019 car ownership and rail passenger trip data to avoid the impact of COVID on the two indicators.

scores and each other) prepared SA scores for each token in this study's datasets. Table 8 presents the correlation between the two human analysts and the machine analysis, showing strong correlations (>0.70) both between the human analysts and with BERT, which tends to lend credibility to the SA results generated from BERT.

Fig. 1 depicts the changes between 2010 and 2020 in sentiment for each category of tokens, as well as the entire documents (as a comparator with respect to the tokens). For simplicity, we averaged the sentiment scores estimated by the two human analysts into the single "human analyst" panel of Fig. 1.

Table 8
Correlation matrix comparing sentiment outputs from BERT and human analysts.

	BERT	Human analyst 1	Human analyst 2
BERT	1.0	+0.77	+0.73
Human analyst 1		1.0	+0.76
Human analyst 2			1.0

Note: All correlations are significant at the p < 0.01 level.

Overall, the sentiment of FYP-T documents as well as the identified tokens increased between 2010 and 2020. Notable findings are: (1) 'Congestion' and 'car' remain the most negative tokens in both 2010 and 2020; (2) 'Autonomous vehicle' was not mentioned in 2010 FYP-Ts, but it is one of the most positive topics in 2020; (3) 'Air pollution' becomes much more negative in 2020, and (4) 'Equity' is negative in 2010 but becomes much more positive in 2020.

Among the tokens in the 'travel option' category, most (except for 'car', as noted above) ranked in the mid-range in terms of sentiment. The positivity of references to 'bike' and 'walking' increased, as we had anticipated, however sentiment towards 'shared mobility' dropped marginally during this time period. 'Public transport' was by far the highest frequency of the modal tokens, and sentiment towards it was consistently (in both 2010 and 2020) higher than 'car' but below-

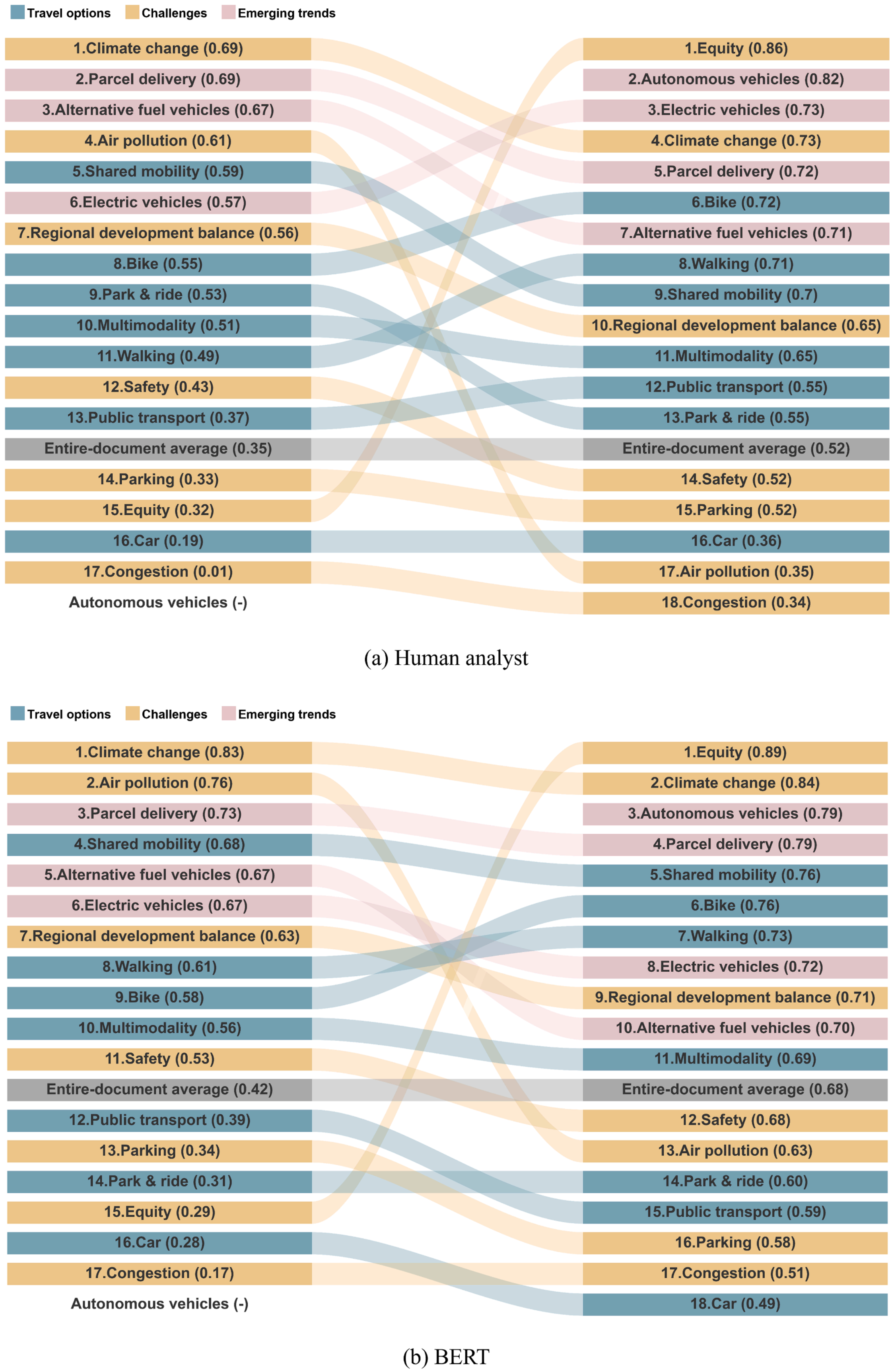

Fig. 1. Changes of rankings for average sentiment scores from 2010 (left column) to 2020 (right column).
Note: There are no references to 'autonomous vehicles' in the 2010 FYP-Ts.

median relative to all tokens. 'Park & ride' has the lowest frequency in this category, and its change in sentiment is unclear.

For the 'challenges' category of references, the converging trend in the sentiment of 'equity' (which increased) and 'regional development balance' (which decreased) are in accordance with their changes of frequency, suggesting a growing focus on both inequity between rural and urban locations as well as disadvantaged social groups.

For the 'emerging trends' category, most topics rank highly in sentiment, but rankings generally decreased between 2010 and 2020. We had expected an increasingly positive trend over time in sentiment towards 'electric vehicles', but the result is mixed across human analysts and BERT. As noted above, although 'autonomous vehicle' appears only in the 2020 documents, it is the most positive of the 'emerging trends' topics in 2020. This indicates generally positive (but as yet untested) expectations for this technology that has not yet been commercially deployed at scale.

5.5. Sentiment and city characteristics correlations

For reasons outlined in Section 5.3 (negative externalities associated with cars are likely to be more serious in larger and more economically developed cities), we expected that sentiment (particularly in 2020, after 10 years of rapid growth in car ownership) towards tokens associated with sustainable travel options and emerging trends will be higher in the high-tier cities. Likewise, we wished to evaluate whether sentiment towards the sustainable travel and emerging trends tokens will associate positively with the population, GRP per capita, car ownership, and rail-passenger journeys city-level indicators.

Table 9 shows that the sentiment scores of the tokens analyzed by city tiers are in general consistent for both human analysts and BERT. In the first-tier cities, sentiment towards tokens associated with sustainable travel options were in some cases below average in 2010, but all are above average in 2020. Interestingly, most sentiment scores of 'travel option' tokens in the second-tier cities are below average according to both human analysts and BERT. Sentiment towards the 'travel option' topics in the third-tier cities are very positive in 2010, however the sentiment scores of 'public transport' and 'multimodality' decrease to become below average in 2020.

Sentiment towards 'challenges' shows generally consistent results between human analysts and BERT for the new-first tier and second-tier cities, but discrepancies arise for the other two tiers, particularly in 2020: In the first-tier cities, human analysts' results show above-average sentiment towards 'congestion' and 'parking', while sentiment towards 'air pollution' and 'climate change' is below average. However, BERT's results show below-average sentiment towards 'congestion' and 'parking', and above-average sentiment towards 'air pollution' and 'climate change'; In the third-tier cities, human analysts' results show above-average sentiment towards 'climate change' and 'air pollution', and below-average sentiment towards 'parking'. BERT's results show below-average sentiment towards 'climate change' and 'air pollution', and above-average sentiment towards 'parking'. The sentiment score of 'equity' in the first-tier city is consistently above average in 2020, which aligns with its high frequencies.

The results for the 'emerging trends' category align well with our a priori expectations. All tokens within the emerging trends category exhibit above-average sentiment scores in the first-tier cities and below-average sentiment scores in the third-tier cities in 2020 (except for the sentiment towards 'autonomous vehicle' identified by BERT). The results for the other two tiers are mixed, but 'parcel delivery' in the new-first and second-tier cities received above-average sentiment scores from both human analysts and BERT in 2020.

Table 10 and Table 11 present the correlations between sentiment scores (as assessed by human analysts and BERT, respectively) and the four city characteristics indicators (population, GRP/capita, car

Table 9
Average sentiment scores of tokens for different tiers.

	Human analysts					**BERT**				
	First tier	**New first tier**	**Second tier**	**Third tier**	**Average**	**First tier**	**New first tier**	**Second tier**	**Third tier**	**Average**
	2010/2020	**2010/2020**	**2010/2020**	**2010/2020**	**2010/2020**	**2010/2020**	**2010/2020**	**2010/2020**	**2010/2020**	**2010/2020**
Travel options										
Car	*0.04*/**0.38**	**0.31**/*0.33*	*0.13*/**0.36**	**0.38/0.38**	0.19/0.36	**0.35/0.61**	**0.31**/*0.45*	*0.17/0.36*	*0.13/0.25*	0.28/0.49
Shared mobility	*0.56*/**0.71**	**0.64/0.74**	*0.50/0.60*	**1.00/0.77**	0.59/0.70	*0.60*/**0.82**	**0.86/0.81**	**1.00**/*0.65*	**1.00/0.79**	0.68/0.76
Bike	*0.50*/**0.76**	**0.59**/*0.70*	*0.20/0.55*	**1.00/0.75**	0.55/0.72	**0.69/0.83**	*0.56/0.73*	*0.20/0.47*	**1.00/0.79**	0.58/0.76
Walking	*0.48*/**0.71**	**0.58/0.71**	*0.14/0.55*	**0.50/0.90**	0.49/0.71	**0.70/0.76**	**0.64/0.73**	*0.14/0.48*	**1.00/0.87**	0.61/0.73
Public transport	**0.45/0.71**	*0.33*/**0.55**	*0.36/0.44*	**0.37**/*0.52*	0.37/0.55	**0.47/0.75**	*0.37*/**0.60**	*0.33/0.48*	**0.39**/*0.56*	0.39/0.59
Multimodality	**0.59/0.77**	*0.36/0.62*	**0.55**/*0.57*	**0.74**/*0.59*	0.51/0.65	**0.69/0.79**	*0.47/0.69*	*0.50/0.59*	**0.69**/*0.67*	0.56/0.69
Park & ride	*0.50/0.33*	**0.58/0.80**	*0.50*/**1.00**	-/**1.00**	0.53/0.55	*0.25/0.33*	**0.50/1.00**	**0.00/1.00**	-/**1.00**	0.31/0.60
Challenges										
Climate change	*0.61/0.70*	*0.61/0.64*	**0.86/0.94**	**0.90/0.79**	0.69/0.73	*0.67*/**0.84**	*0.81/0.83*	**1.00/0.94**	**1.00**/*0.79*	0.83/0.84
Air pollution	*0.18/0.27*	**0.75**/*0.27*	*0.61*/**0.44**	**0.63/0.45**	0.61/0.35	*0.36*/**0.67**	**0.78**/*0.54*	**0.86/0.78**	**0.95**/*0.58*	0.76/0.63
Congestion	**0.24/0.37**	*−0.05/0.20*	*−0.08*/**0.42**	*−0.63*/**0.63**	0.01/0.34	**0.35**/*0.49*	*0.16/0.35*	*0.00*/**0.75**	*−0.25*/**0.63**	0.17/0.51
Parking	**0.52/0.54**	*0.14*/**0.60**	*0.06/0.39*	**0.46**/*0.50*	0.33/0.52	**0.52**/*0.58*	*0.20*/**0.65**	*0.00/0.53*	**0.50/0.60**	0.34/0.58
Safety	**0.50**/*0.40*	*0.32/0.44*	**0.44**/*0.50*	**0.50/0.73**	0.43/0.52	**0.58**/*0.62*	*0.36/0.62*	**0.69/0.79**	**0.58/0.73**	0.53/0.68
Equity	*0.50*/**0.97**	*0.50*/**0.89**	**0.75**/*0.83*	-/*0.77*	0.57/0.86	**0.50/1.00**	**0.00/0.78**	*0.50*/**1.00**	-/*0.82*	0.43/0.89
Regional development balance	*0.51*/**0.66**	**0.62/0.69**	*0.41*/**0.72**	**0.58**/*0.59*	0.54/0.65	*0.53*/**0.81**	**0.72/0.72**	*0.45*/**0.79**	**0.64**/*0.62*	0.60/0.71
Emerging trends										
Alternative fuel vehicle	*0.65*/**0.75**	**0.71**/*0.68*	**0.81/0.76**	*0.62/0.65*	0.67/0.71	**0.77/0.74**	**0.83**/*0.67*	**0.88/0.78**	*0.52/0.61*	0.67/0.70
Electric vehicles	**0.57/0.86**	**0.75**/*0.56*	*0.38/0.73*	*0.25/0.70*	0.57/0.73	**0.79/0.87**	**0.75**/*0.54*	*0.50*/**0.74**	*0.00/0.63*	0.67/0.72
Autonomous vehicles	-/**0.85**	-/**0.83**	-/*0.75*	-/*0.75*	-/0.82	-/**0.95**	-/*0.67*	-/**0.83**	-/**0.83**	-/0.79
Parcel delivery	*0.60*/**0.79**	*0.61*/**0.73**	*0.67*/**0.74**	**0.76**/*0.67*	0.69/0.72	*0.60*/**0.85**	*0.67*/**0.83**	**0.78/0.81**	**0.78**/*0.71*	0.73/0.79

Note 1: The format X/Y denotes token sentiment scores in years 2010 (X) and 2020 (Y).
Note 2: "-" represents that there are no references of the tokens in the FYP-Ts.
Note 3: Bold and italics indicate the average sentiment score of a certain token is higher and lower than the average of all cities, respectively.

Table 10
Correlation between average sentiment scores (human) of tokens and characteristics of the cities.

	Population (1,000 people)		GRP per capita (CNY)		Car ownership per 1,000 people		Rail passenger trips per year (1,000 people)	
	2010	2020	2010	2020	2010	2020	2010	2020
Travel options								
Bike								
Car								
Multimodality				0.37 (0.08)				
Park & ride								
Shared mobility			−0.73 (0.06)					
Public transport		0.36 (0.09)		*0.45 (0.03)*				
Walking						−0.41 (0.07)		
Challenges								
Air pollution								
Climate change	−0.50 (0.05)	−0.40 (0.06)						
Congestion								
Equity								
Parking								
Regional development balance							0.45 (0.05)	
Safety								−0.39 (0.09)
Emerging trends								
Alternative fuel vehicle						0.36 (0.09)		
Autonomous vehicles	–		–		–		–	**−0.88 (<0.01)**
Electric vehicles								
Parcel delivery								

Note 1: Correlations with *p*-value > 0.10 were excluded in the table.
Note 2: There are no references to 'autonomous vehicles' in the 2010 FYP-Ts.
Note 3: Bold and italics indicate significance at the 0.01 and 0.05 levels, respectively.
Note 4: We used 2019 car ownership and rail passenger trip data to avoid the impact of COVID on the two indicators.

Table 11
Correlation between average sentiment scores (BERT) of tokens and characteristics of the cities.

	Population (1,000 people)		GRP per capita (CNY)		Car ownership per 1,000 people		Rail passenger trips per year (1,000 people)	
	2010	2020	2010	2020	2010	2020	2010	2020
Travel options								
Bike								
Car								
Multimodality								
Park & ride							0.64 (0.09)	
Shared mobility			*−0.84 (0.02)*		−0.77 (0.07)			
Public transport		0.36 (0.09)		**0.54 (<0.01)**				
Walking			0.46 (0.10)					
Challenges								
Air pollution	**−0.74 (<0.01)**		*−0.52 (0.04)*					
Climate change	**−0.64 (<0.01)**							
Congestion								
Equity								
Parking						−0.41 (0.08)		
Regional development balance						*0.45 (0.03)*		
Safety								
Emerging trends								
Alternative fuel vehicle						0.40 (0.06)		
Autonomous vehicles	–		–		–		–	*−0.65 (0.01)*
Electric vehicles							0.50 (0.08)	
Parcel delivery						**0.54 (<0.01)**		

Note 1: Correlations with *p*-value > 0.10 were excluded in the table.
Note 2: There are no references to 'autonomous vehicles' in the 2010 FYP-Ts.
Note 3: Bold and italics indicate significance at the 0.01 and 0.05 levels, respectively.
Note 4: We used 2019 car ownership and rail passenger trip data to avoid the impact of COVID on the two indicators.

ownership, rail passenger journeys). Comparing with Table 7 which shows correlations between token frequencies and the indicators, these two tables (which have results that are not statistically significant) are very sparse, and there are few consistent findings observed across the results of human analysts and BERT (correlations between 'climate change' and population in 2010, 'public transport' and population/GRP per capita in 2020, 'shared mobility' and GRP per capita in 2010, 'alternative fuel vehicle' and car ownership in 2020, and 'autonomous vehicle' and rail passenger trips in 2020). The results for 'climate change', 'public transport' and 'alternative fuel vehicle' are as expected, and we note the observed strong negative correlation between 'autonomous vehicles' and rail passenger trips, found both by the BERT algorithm and human analysts.

6. Conclusions

Urban transport policy in Chinese cities has a wide-ranging set of impacts both locally and globally. This paper aims to document patterns in the structure and content of strategic transport planning documents (Five Year Plans for Transport) of Chinese cities', motivated to understand how the documents vary across different types of cities, as well as over time between 2010 and 2020. A brief summary of findings follows, followed by a discussion of further research needs.

First, we find that Chinese cities' FYP-Ts have tended to lengthen during the 2010s, and are relatively text-heavy documents with indicators of difficult readability. Their writing style was found to be relatively complex, and became increasingly complex during the period 2010 – 2020. It could be useful for the Chinese government to consider initiatives similar to the U.S.'s Plain Writing Act (Office of the Federal Register, 2010) to enhance the accessibility of the content of FYP-T documents to members of the general public with average reading abilities.

Second, we document systematic differences in the content of FYP-Ts on the basis of the level of economic development. More economically advanced cities focused greater attention with their FYP-Ts on sustainable travel options (such as bike, walking, and public transport), negative externalities related to cars (such as congestion, climate change, and parking), as well as emerging technologies (such as autonomous vehicles and electric vehicles). In contrast, less economically developed cities devote heightened attention to regional development issues (like revenue-sharing, spatial equity in transport network investment, and general economic development objectives). While there are some associations between the focus of the policymakers (proxied by token frequency) and city-level transport indicators, these associations are much weaker in comparison to the geodemographic indicators.

Third, using a sentiment analysis approach with an algorithm trained to the specific content of FYP-Ts (as well as human analysis for comparison purposes), we find that the general sentiment of FYP-Ts became increasingly positive during the period 2010 – 2020. With respect to specific topics, the two largest changes are sentiment towards 'equity' becoming much more positive, and vice versa for 'air pollution'. In contrast to this finding regarding 'air pollution', sentiment towards 'congestion' is consistently very negative in both 2010 and 2020. These results (and all noted in this listing) relate to sentiment hold for both the SA algorithm's and human analysts' findings.

Fourth, in terms of specific modes of transport, the frequency at which 'cars' are mentioned within FYP-Ts decreased between 2010 and 2020, and sentiment within the FYP-Ts towards 'cars' is strongly negative in both years. In comparison to automobiles, 'public transport' is mentioned more frequently and references to it are more positive, however references to 'walking' and 'bicycling' are much more positive (and consistently so across both 2010 and 2020). With regards to the associations between sentiment scores and city characteristics, the sentiment towards the tokens generally correlates weakly with the selected city geodemographic and transport indicators.

Fifth and finally, 'autonomous vehicles' are not referenced in any of the FYP-Ts in 2010, and when this topic appears in 2020, it is one of the topics with the most strongly positive sentiment. Sentiment towards non-internal-combustion propulsion technologies ('electric vehicles' and 'alternative fuel vehicles' in general) is strongly positive in both years, with increasing frequency of mentions with the FYP-Ts. The same was also found with respect to 'parcel delivery' and 'equity', with the former experienced the greatest increase in frequency-of-references of all topics considered and the latter experienced a shape increase in both token frequency and sentiment scores in this study.

We conclude this paper with a brief discussion of future research needs.

Expanding the scope of the current analysis to include more of the FYP-Ts from China's 687 cities and additional points in time would be valuable to either confirm or reject the findings from this small sample. This, however, would be a non-trivial task given the labor-intensive nature of this line of research. A strategy for future research could therefore include machine learning techniques to automate portions of the workflow that were performed manually in this study. This could allow trends in transport planning (in China as well as elsewhere) to be identified in real time, providing enhanced situational awareness across the profession.

A standard approach that could automate topic/theme identification would also be useful, by providing researchers the ability to publish summaries of emerging transport planning trends in FYP-Ts and similar documents in near-real-time with less labor effort than in the present study. This would provide a valuable resource for practitioners in any given city to better understand the wider context in which they are working.

Future research is also needed to establish how the patterns and trends identified in this study of Chinese cities compare to urban transport planning documents in different societies. International comparisons and benchmarking may yield greater insight into the practice of transport planning in different contexts.

It is hoped that this line of research will assist researchers and policymakers in understanding how the emphasis placed on transport planning topics varies across Chinese cities, as well as how this has evolved in recent years.

CRediT authorship contribution statement

Meng Guo: Conceptualization, Data curation, Methodology, Writing – original draft. **Qingyang Li:** Data curation, Methodology, Writing – original draft. **Chenyang Wu:** Conceptualization, Data curation, Methodology, Investigation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Scott Le Vine:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Gang Ren:** Data curation, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 52102389) and the Northwestern Polytechnical University Start-up funding (D5000230159).

References

Association of Bay Area Governments, Metropolitan Transportation Commission, 2021. Plan Bay Area 2050: A vision for the future [WWW Document]. URL https://www.planbayarea.org/sites/default/files/documents/Plan_Bay_Area_2050_October_2021.pdf (accessed 8.1.22).

Baj-Rogowska, A., 2017. Sentiment analysis of Facebook posts: The Uber case. In: 2017 IEEE 8th International Conference on Intelligent Computing and Information Systems, ICICIS 2017. Institute of Electrical and Electronics Engineers Inc., pp. 391–395. https://doi.org/10.1109/INTELCIS.2017.8260068

Bengtsson, M., 2016. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2, 8–14. https://doi.org/10.1016/j.npls.2016.01.001.

Bista, S., Hollander, J.B., Situ, M., 2021. A content analysis of transportation planning documents in Toronto and Montreal. Case Stud. Transp. Policy 9, 1–11. https://doi.org/10.1016/j.cstp.2020.06.007.

Boisjoly, G., El-Geneidy, A.M., 2017. How to get there? A critical assessment of accessibility objectives and indicators in metropolitan transportation plans. Transp. Policy 55, 38–50. https://doi.org/10.1016/j.tranpol.2016.12.011.

Casas, I., Delmelle, E.C., 2017. Tweeting about public transit — Gleaning public perceptions from a social media microblog. Case Stud. Transp. Policy 5, 634–642. https://doi.org/10.1016/j.cstp.2017.08.004.

Casey, J., Koleski, K., 2011. Backgrounder: China's 12th Five-Year Plan.

Chan, H.-Y., Cheung, K.K.C., Erduran, S., 2023. Science communication in the media and human mobility during the COVID-19 pandemic: a time series and content analysis. Public Health 218, 106–113. https://doi.org/10.1016/j.puhe.2023.03.001.
Chen, Y., Woo, R., 2020. China's home price growth at almost 1-1/2-year low, further softening seen | Reuters [WWW Document]. Reuters. URL https://www.reuters.com/article/uk-china-economy-houseprices-idUKKBN1ZF054 (accessed 5.31.23).
New York Metropolitan Transportation Council, 2021. Moving Forward: Your Region, Connected. [WWW Document]. URL https://www.nymtc.org/movingforward/the-complete-plan/index.html (accessed 8.1.22).
Cui, Y., Che, W., Liu, T., Qin, B., Yang, Z., 2021. Pre-Training with Whole Word Masking for Chinese BERT. IEEE/ACM Trans. Audio Speech Lang. Process. 29, 3504–3514. https://doi.org/10.1109/TASLP.2021.3124365.
Da Silva, N.F.F., Hruschka, E.R., Hruschka, E.R., 2014. Tweet sentiment analysis with classifier ensembles. Decis. Support Syst. 66, 170–179. https://doi.org/10.1016/j.dss.2014.07.003.
Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf. 1, 4171–4186. doi: 10.48550/arxiv.1810.04805.
Ecola, L., Rohr, C., Zmud, J., Kuhnimhof, T., Phleps, P., 2019. Driving in the Future in Developing Countries. Driving in the Future in Developing Countries. https://doi.org/10.7249/rb9794.
Elvy, J., 2014. Public participation in transport planning amongst the socially excluded: An analysis of 3rd generation local transport plans. Case Stud. Transp. Policy 2, 41–49. https://doi.org/10.1016/j.cstp.2014.06.004.
Fan, C.C., 2006. China's Eleventh Five-Year Plan (2006–2010): From "Getting Rich First" to "Common Prosperity". Eurasian Geography and Economics 47, 708–723. https://doi.org/10.2747/1538-7216.47.6.708.
Fang, X., Zhan, J., 2015. Sentiment analysis using product review data. J. Big Data 2, 5. https://doi.org/10.1186/s40537-015-0015-2.
Gitto, S., Mancuso, P., 2017. Improving airport services using sentiment analysis of the websites. Tour. Manag. Perspect. 22, 132–136. https://doi.org/10.1016/j.tmp.2017.03.008.
Yicai Global, 2022. Why are they the new first-tier cities of 2022? (in Chinese) [WWW Document]. URL https://www.yicai.com/news/101430657.html (accessed 7.31.22).
Heikkila, E., Xu, Y., 2014. Seven Prototypical Chinese Cities. Urban Stud. 51, 827–847. https://doi.org/10.1177/0042098013492231.
Heilmann, S., 2011. China Analysis 92 October 2011 Making Plans for Markets : Policy for the Long Term in China. China Anal. 13, 33–40.
Hu, A.G., 2016. The Five-Year Plan: A new tool for energy saving and emissions reduction in China. Adv. Clim. Chang. Res. 7, 222–228. https://doi.org/10.1016/j.accre.2016.12.005.
Huang, S., Wang, D., 2019. Construction, performance and application of New Era People's Daily Segmented Corpus (III): Analysis and comparison of sentence length and word. Libr. Inf. Serv. 63, 5–15. https://doi.org/10.13266/j.issn.0252-3116.2019.24.001 in Chinese.
Jamil, M.L., Pais, S., Cordeiro, J., Dias, G., 2022. Detection of extreme sentiments on social networks with BERT. Social Network Analysis and Mining 12, 1–16. https://doi.org/10.1007/S13278-022-00882-Z/FIGURES/10.
Jia, M., Liu, Y., Lieske, S.N., Chen, T., 2020. Public policy change and its impact on urban expansion: An evaluation of 265 cities in China. Land Use Policy 97, 104754.
Kang, D., Park, Y., 2014. Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach. Expert Syst. Appl. 41, 1041–1050. https://doi.org/10.1016/j.eswa.2013.07.101.
Kauffmann, E., Peral, J., Gil, D., Ferrández, A., Sellers, R., Mora, H., 2020. A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making. Ind. Mark. Manag. 90, 523–537. https://doi.org/10.1016/j.indmarman.2019.08.003.
Kaya, M., Fidan, G., Toroslu, I.H., 2012. Sentiment Analysis of Turkish Political News. In: 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology. IEEE, Macau, China, pp. 174–180. https://doi.org/10.1109/WI-IAT.2012.115.
Kouloumpis, E., Wilson, T., Moore, J., 2021. Twitter Sentiment Analysis: The Good the Bad and the OMG! Proc. Int. AAAI Conf. Web Soc. Media 5, 538–541. https://doi.org/10.1609/icwsm.v5i1.14185.
Kumar, C.H., Malathi, K., 2022. Analysing the Sentiments in e-newspaper Contents using Novel Bidirectional Encoder Representation for Transformation - BERT over Linear Regression Algorithm. J. Pharm. Negat. Results 13, 30–38. https://doi.org/10.47750/PNR.2022.13.S04.004.
Kumar, S., Yadava, M., Roy, P.P., 2019. Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. Inf. Fusion 52, 41–52. https://doi.org/10.1016/j.inffus.2018.11.001.
Kwon, H.-J., Ban, H.-J., Jun, J.-K., Kim, H.-S., 2021. Topic Modeling and Sentiment Analysis of Online Review for Airlines. Information 12, 78. https://doi.org/10.3390/info12020078.
Lee, R.J., Sener, I.N., 2016. Transportation planning and quality of life: Where do they intersect? Transp Policy 48, 146–155. https://doi.org/10.1016/j.tranpol.2016.03.004.
Lesteven, G., Thébert, M., 2022. Who cares about AVs? Insights from French media discourse on Twitter and in the press. Case Stud. Transp. Policy 10, 1078–1089. https://doi.org/10.1016/j.cstp.2022.03.022.
Li, H., Dong, X., Jiang, Q., Dong, K., 2021. Policy analysis for high-speed rail in China: Evolution, evaluation, and expectation. Transp. Policy 106, 37–53. https://doi.org/10.1016/j.tranpol.2021.03.019.
Liang, X.H., Tan, K.H., 2019. Market potential and approaches of parcels and mail by high speed rail in China. Case Stud. Transp. Policy 7, 583–597. https://doi.org/10.1016/j.cstp.2019.07.008.
Ma, W., de Jong, M., Hoppe, T., de Bruijne, M., 2021. From city promotion via city marketing to city branding: Examining urban strategies in 23 Chinese cities. Cities 116, 103269.
Manaugh, K., Badami, M.G., El-Geneidy, A.M., 2015. Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in north america. Transp. Policy 37, 167–176. https://doi.org/10.1016/j.tranpol.2014.09.013.
Mäntylä, M.V., Graziotin, D., Kuutila, M., 2018. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. Comput. Sci. Rev. 27, 16–32. https://doi.org/10.1016/j.cosrev.2017.10.002.
Martin-Domingo, L., Martín, J.C., Mandsberg, G., 2019. Social media as a resource for sentiment analysis of Airport Service Quality (ASQ). J. Air Transp. Manag. 78, 106–115. https://doi.org/10.1016/j.jairtraman.2019.01.004.
Medhat, W., Hassan, A., Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams Eng. J. 5, 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011.
Miami-Dade Transportation Planning Organization, 2019. Miami-Dade 2045 LRTP [WWW Document]. URL https://miamidade2045lrtp.com/the-plan (accessed 8.1.22).
Moody, J., Wang, S., Chun, J., Ni, X., Zhao, J., 2019. Transportation policy profiles of Chinese city clusters: A mixed methods approach. Transp. Res. Interdiscip. Perspect. 2, 100053 https://doi.org/10.1016/j.trip.2019.100053.
Morshed, S.A., Khan, S.S., Tanvir, R.B., Nur, S., 2021. Impact of The Covid-19 Pandemic on Ride-Hailing Services Based on Large-Scale Twitter Data Analysis. J. Urban Manag. https://doi.org/10.1016/j.jum.2021.03.002.
NBS, 2011. Method of dividing China into East, West, Central, and Northeast regions. (in Chinese) [WWW Document]. URL http://www.stats.gov.cn/ztjc/zthd/sjtjr/dejtjkfr/tjkp/201106/t20110613_71947.htm (accessed 7.31.22).
NBS, 2022. China Statistical Yearbook [WWW Document]. URL http://www.stats.gov.cn/tjsj/ndsj/2022/indexch.htm.
Office of the Federal Register, N.A. and R.A., 2010. Public Law 111 - 274 - Plain Writing Act of 2010. govinfo.gov.
Pagolu, V.S., Reddy, K.N., Panda, G., Majhi, B., 2017. Sentiment analysis of Twitter data for predicting stock market movements, in: International Conference on Signal Processing, Communication, Power and Embedded System, SCOPES 2016 - Proceedings. IEEE, Paralakhemundi, Odisha, India, pp. 1345–1350. doi: 10.1109/SCOPES.2016.7955659.
Park, J.Y., Mistur, E., Kim, D., Mo, Y., Hoefer, R., 2022. Toward human-centric urban infrastructure: Text mining for social media data to identify the public perception of COVID-19 policy in transportation hubs. Sustain. Cities Soc. 76, 103524 https://doi.org/10.1016/j.scs.2021.103524.
Qi, B., Costin, A., Jia, M., 2020. A framework with efficient extraction and analysis of Twitter data for evaluating public opinions on transportation services. Travel Behav. Soc. 21, 10–23. https://doi.org/10.1016/j.tbs.2020.05.005.
Rao, Y., Lei, J., Wenyin, L., Li, Q., Chen, M., 2014. Building emotional dictionary for sentiment analysis of online news. World Wide Web 17, 723–742. https://doi.org/10.1007/s11280-013-0221-9.
Ratanaburi, N., Alade, T., Saçli, F., 2021. Effects of stakeholder participation on the quality of bicycle infrastructure. A case of Rattanakosin bicycle lane, Bangkok, Thailand. Case Stud. Transp. Policy 9, 637–650. https://doi.org/10.1016/j.cstp.2021.03.002.
Rill, S., Reinel, D., Scheidt, J., Zicari, R.V., 2014. PoliTwi: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. Knowledge-Based Syst. 69, 24–33. https://doi.org/10.1016/j.knosys.2014.05.008.
Rogoff, K.S., Yang, Y., 2022. A Tale of Tier 3 Cities. doi: 10.3386/W30519.
Sari, I.C., Ruldeviyani, Y., 2020. Sentiment Analysis of the Covid-19 Virus Infection in Indonesian Public Transportation on Twitter Data: A Case Study of Commuter Line Passengers, in: 2020 International Workshop on Big Data and Information Security (IWBIS). IEEE, Depok, Indonesia, pp. 23–28. https://doi.org/10.1109/IWBIS50925.2020.9255531.
Serrano-Guerrero, J., Olivas, J.A., Romero, F.P., Herrera-Viedma, E., 2015. Sentiment analysis: A review and comparative analysis of web services. Inf. Sci. (Ny) 311, 18–38. https://doi.org/10.1016/j.ins.2015.03.040.
Singh, M., Jakhar, A.K., Pandey, S., 2021. Sentiment analysis on the impact of coronavirus in social life using the BERT model. Social Network Analysis and Mining 11, 1–11. https://doi.org/10.1007/S13278-021-00737-Z/FIGURES/12.
Singleton, P.A., Clifton, K.J., 2017. Considering health in US metropolitan long-range transportation plans: A review of guidance statements and performance measures. Transp. Policy 57, 79–89. https://doi.org/10.1016/j.tranpol.2017.02.003.
Souma, W., Vodenska, I., Aoyama, H., 2019. Enhanced news sentiment analysis using deep learning methods. J. Comput. Soc. Sci. 2, 33–46. https://doi.org/10.1007/s42001-019-00035-x.
The State Council of the People's Republic of China, 2021a. Economists Interpret the "14th Five-Year Plan" and the Outline of Long-term Objectives in 2035 (in Chinese) [WWW Document]. URL http://www.gov.cn/zhengce/2021-03/14/content_5592819.htm (accessed 3.7.23).
The State Council of the People's Republic of China, 2021b. The 14th Five-Year Plan (2021–2025) for National Economic and Social Development of the People's Republic of China (in Chinese) [WWW Document]. URL https://www.gov.cn/zhuanti/shisiwu/chrome/index.html#!/index (accessed 3.7.23).
U.S. General Accounting Office, 1996. PEMD-10-3.1, Content Analysis: A Methodology for Structuring and Analyzing Written Material.

Wang, S., Moody, J., Zhao, J., 2021. What prompts the adoption of car restriction policies among Chinese cities. Int. J. Sustain. Transp. 15, 559–570. https://doi.org/10.1080/15568318.2020.1770905.

Wang, B., Wu, C., Kang, L., Reniers, G., Huang, L., 2018. Work safety in China's Thirteenth Five-Year plan period (2016–2020): Current status, new challenges and future tasks. Saf. Sci. 104, 164–178. https://doi.org/10.1016/j.ssci.2018.01.012.

Wen, S., Jia, Z., Chen, X., 2022. Can low-carbon city pilot policies significantly improve carbon emission efficiency? Empirical evidence from China. J. Clean. Prod. 346, 131131 https://doi.org/10.1016/j.jclepro.2022.131131.

Wu, C., Le Vine, S., Bengel, E., Czerwinski, J., Polak, J., 2022. Sentiment analysis of popular-music references to automobiles, 1950s to 2010s. Transportation (Amst). 49, 641–678. https://doi.org/10.1007/s11116-021-10189-1.

Xinhua Net, 2021. A grand blueprint that gathers the power of hundreds of millions of people——The "14th Five-Year Plan" and the 2035 long-term goal outline compilation record (in Chinese) [WWW Document]. URL http://www.xinhuanet.com/politics/2021-03/19/c_1127231818.htm (accessed 3.7.23).

Xu, Y., Heikkila, E.J., 2020. How can cities learn from each other? Evidence from China's five-year plans. J. Urban Manag. 9, 216–227. https://doi.org/10.1016/j.jum.2020.04.002.

Yadav, A., Vishwakarma, D.K., 2020. Sentiment analysis using deep learning architectures: a review. Artif. Intell. Rev. 53, 4335–4385. https://doi.org/10.1007/s10462-019-09794-5.

Yang, L., Li, Y., Wang, J., Sherratt, R.S., 2020. Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. IEEE Access 8, 23522–23530. https://doi.org/10.1109/ACCESS.2020.2969854.

Yavetz, G., Aharony, N., 2023. Information under lockdown: A content analysis of government communication strategies on Facebook during the COVID-19 outbreak. Heliyon 9 (4), e15562.

Yu, L., Xue, B., Stückrad, S., Thomas, H., Cai, G., 2020. Indicators for energy transition targets in China and Germany: A text analysis. Ecol. Ind. 111, 106012 https://doi.org/10.1016/J.ECOLIND.2019.106012.

Yuan, X., Zuo, J., 2011. Transition to low carbon energy policies in China-from the Five-Year Plan perspective. Energy Policy 39, 3855–3859. https://doi.org/10.1016/j.enpol.2011.04.017.

Zha, W., Ye, Q., Li, J., Ozbay, K., 2023. A social media Data-Driven analysis for transport policy response to the COVID-19 pandemic outbreak in Wuhan, China. Transp. Res. Part A Policy Pract. 172, 103669 https://doi.org/10.1016/j.tra.2023.103669.

Zhang, X., Bai, X., 2017. Incentive policies from 2006 to 2016 and new energy vehicle adoption in 2010–2020 in China. Renew. Sustain. Energy Rev. 70, 24–43. https://doi.org/10.1016/j.rser.2016.11.211.