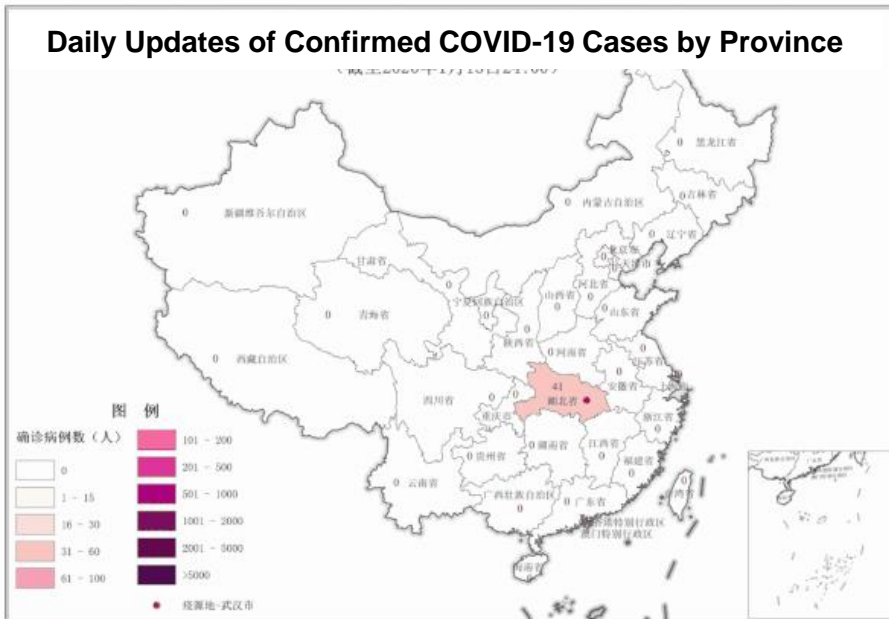


# An Introduction to the Data Resources for COVID-19 Studies

Shuming Bao  
China Data Institute

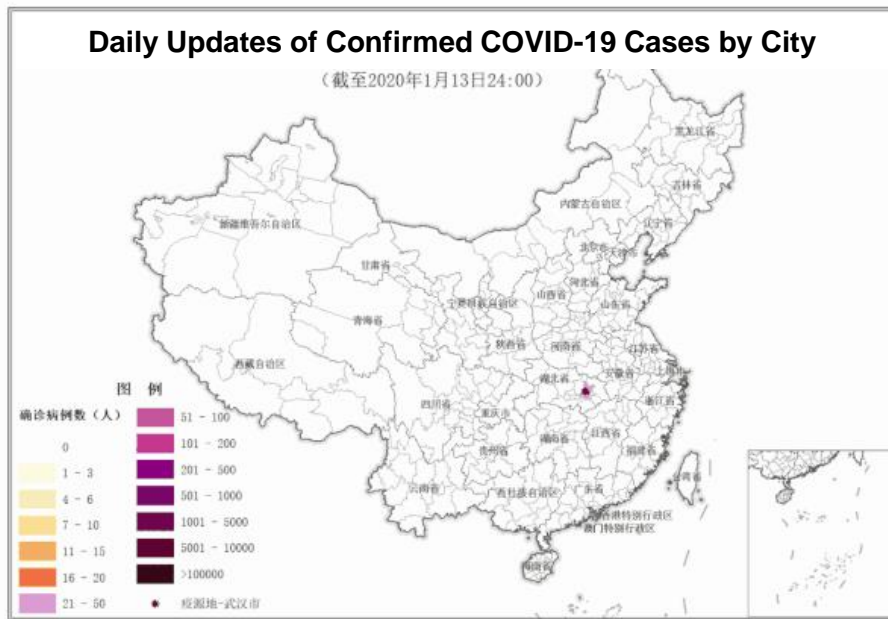
# Daily Updates of Confirmed COVID-19 Cases in China

## Daily Updates of Confirmed COVID-19 Cases by Province



制作单位：中国科学院地理资源所资源与环境信息系统国家重点实验室  
资料来源：丁香园等互联网平台

## Daily Updates of Confirmed COVID-19 Cases by City

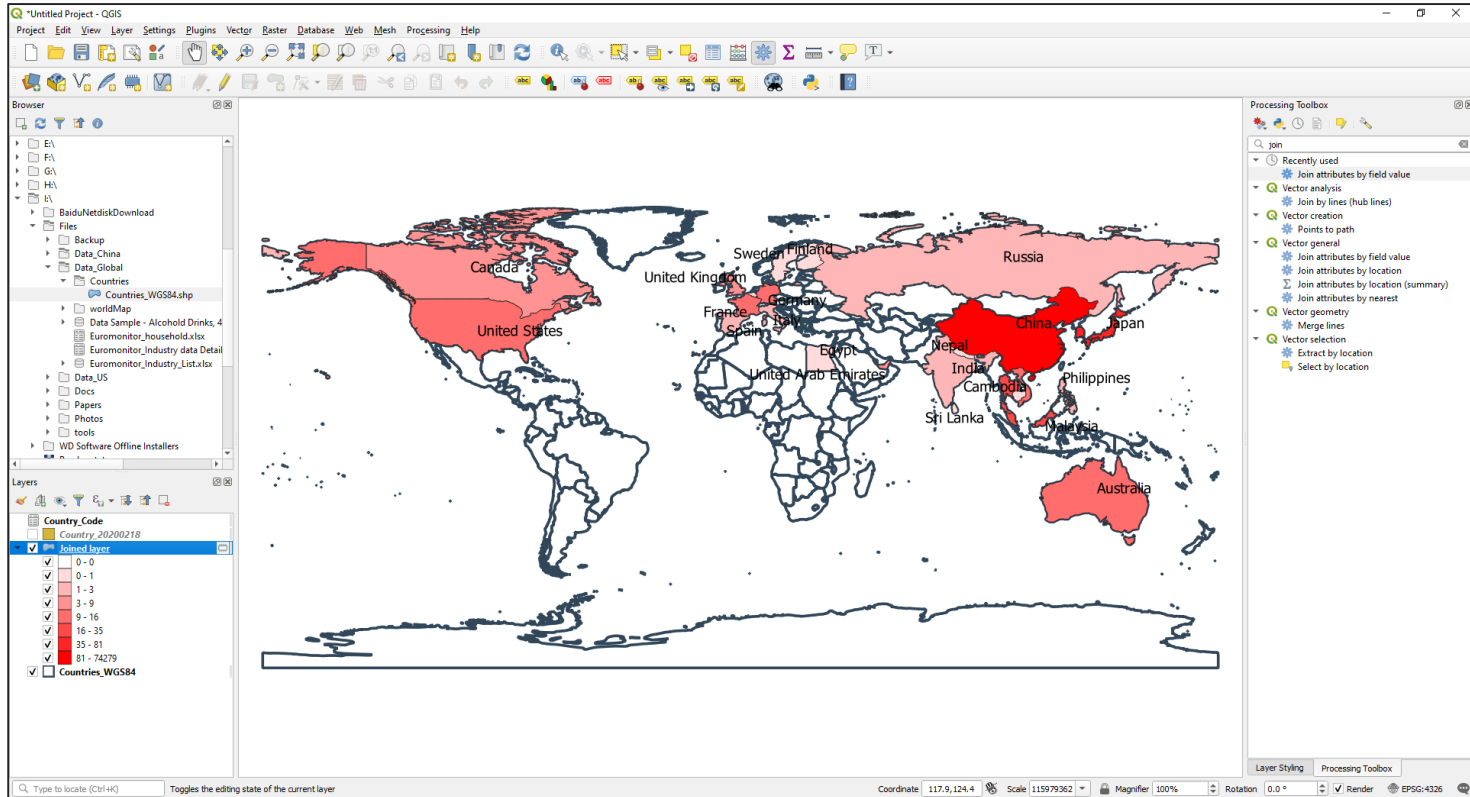


制作单位：中国科学院地理资源所资源与环境信息系统国家重点实验室  
资料来源：丁香园等互联网平台

Source: The project team for mapping story, the Institute of Geographical Science and Resources of the Chinese Academy of Sciences  
中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室图说工作组

[https://mp.weixin.qq.com/s/tw0VKen0H1vCSAJnZJm1\\_Q](https://mp.weixin.qq.com/s/tw0VKen0H1vCSAJnZJm1_Q)

# The World Map of COVID-19 Cases As of 2/18/2020



**Data source:** Ding Xiang Yuan, Real-time virus report ( <https://ncov.dxy.cn/ncovh5/view/pneumonia> )  
**Mapping tool:** QGIS ( <http://qgis.com> )

# 17年前的SARS详细数据，何时公开？

## Where can we find the data for SARS of 17 years ago?

Source: <https://zhuanlan.zhihu.com/p/107292732>



pixabay.com

撰文 | 夏志坚

责编 | 陈晓雪



截至2月15日24时，中国大陆的新冠肺炎累计确诊病例达到68500例，疑似病例8228例，累计死亡病例达到1665人。

这场始于2019年年底的疫情，令不少人回想起在2003年春季爆发的SARS疫情。两者的发展似乎存在不小的相似性：病原体都是由动物传染到人的冠状病毒，都在上一年的年末发现无法确诊的肺炎病例，都遭遇过春运的人流高峰（2003年的除夕是1月31

在国家卫健委的官方网站，以“SARS”为关键词搜索，718条结果（截至2月16日下午）亦缺乏中国大陆地区疫情各阶段的详细数据 [3]。其中有一条链接名为“SARS防治专题”，但打开链接，里面没有任何与SARS有关的有效信息，所有子链接都是空白

# Goal and Objectives

**Goal:** to provide an information infrastructure for the global study of novel coronavirus (COVID-19)

## **Objectives:**

- ❑ To establish a **permanent collection** for the COVID-19 study at local, regional and global levels with information collected and integrated from different sources
- ❑ To facilitate the quantitative research on spatial spread and impacts of COVID-19 study with advanced **methodology and technology**
- ❑ To promote collaborative research in the COVID-19 study with the **cloud platform**
- ❑ To promote the use and sharing of data sources in **teaching and research** with DataVerse/WorldMap/CDL
- ❑ To build the **capacity** for future collaborative projects

# Teams



CGA

**The Center for Geographical Analysis (CGA) at Harvard University.** Its core mission is to support research and teaching in all disciplines across Harvard University with emerging **geospatial technologies**.



CDI

**The China Data Institute, a Michigan based not-for-profit organization.** It aims to promote the use and sharing of **China data**; support quantitative research on China in **social science, digital humanity** and other research subjects.



GCSS

**The GeoComputation Center for Social Science at Wuhan University.** It promotes the scientific research on the theory and method of spatial data in scientific **research**, personnel **training**, international **cooperation** and social **practice**.

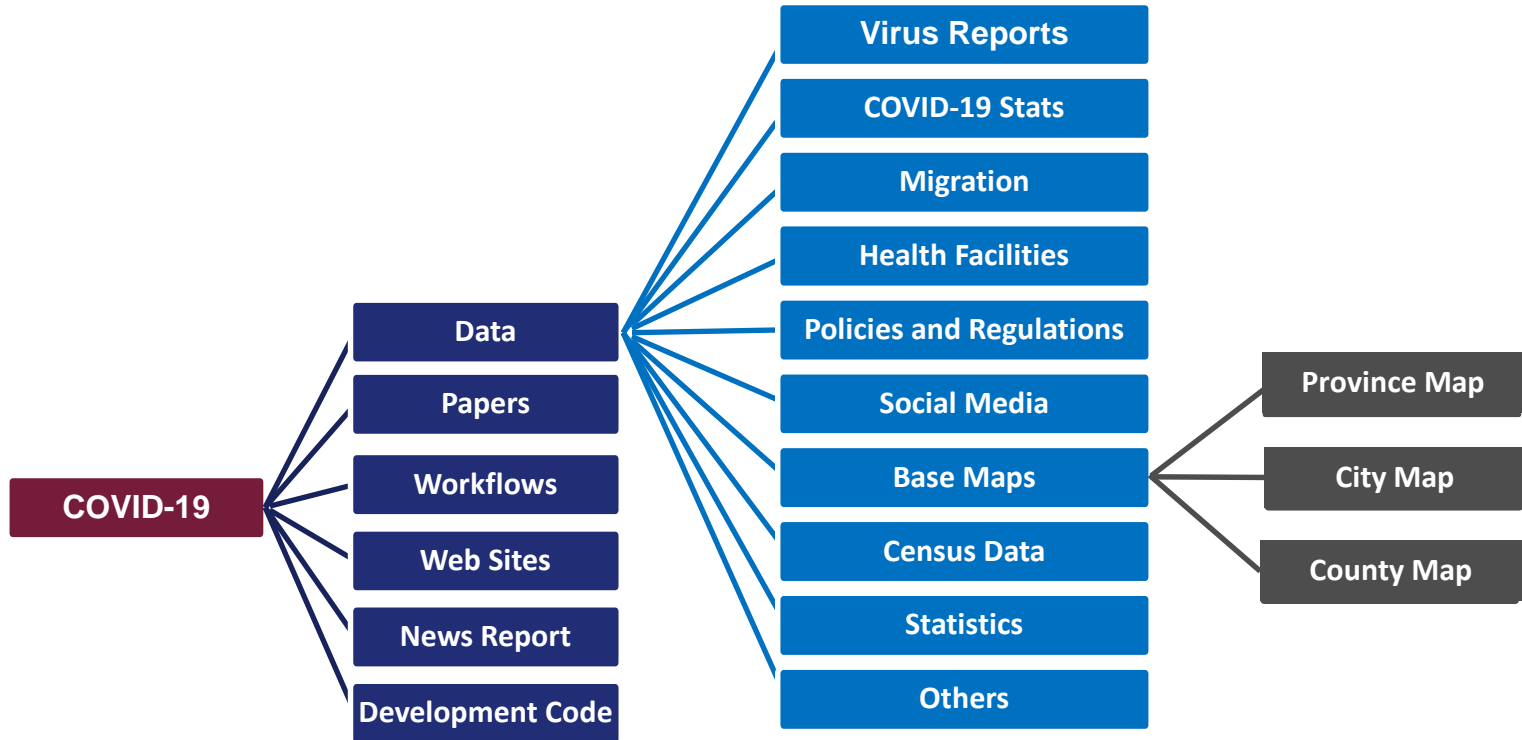


The RMDS Lab is designed to serve a global community of companies and data scientists, empowering them to reap the benefits of **big data**, **artificial intelligence**, and **machine learning**.

# Tasks

- ❑ **Data collection (WHU)**
  - Reported data of virus cases at country, province, and city levels
  - Migration
  - Social media
  - New polices and regulations on virus control by China and other countries
  - Links to other web sites on related projects
  - Related publications
- ❑ **Data integration (WHU and CDI)**
  - Maps
  - Gov't statistics and Census data
- ❑ **Data report (RMDS, WHU, CDI and CGA)**
  - Data analysis
  - PPTs
  - Workflows
- ❑ **Data distribution (CGA)**
  - DataVerse for data management at Harvard (<http://covid-19.chinadatalab.net>)
  - Workflow on Spatial Data Lab at Harvard (<http://harvard.chinadatalab.org>)
  - Project description and events on <http://chinadatalab.net>
- ❑ **Workshops and seminars (CDI, RMDS, WHU and CGA)**
  - China Data Webinar co-hosted by RMDS and China Data Institute
  - Harvard workshop on March 27, 2020, hosted by Harvard Fairbank Center for Chinese Studies

# Data Structure of Resources for COVID-19



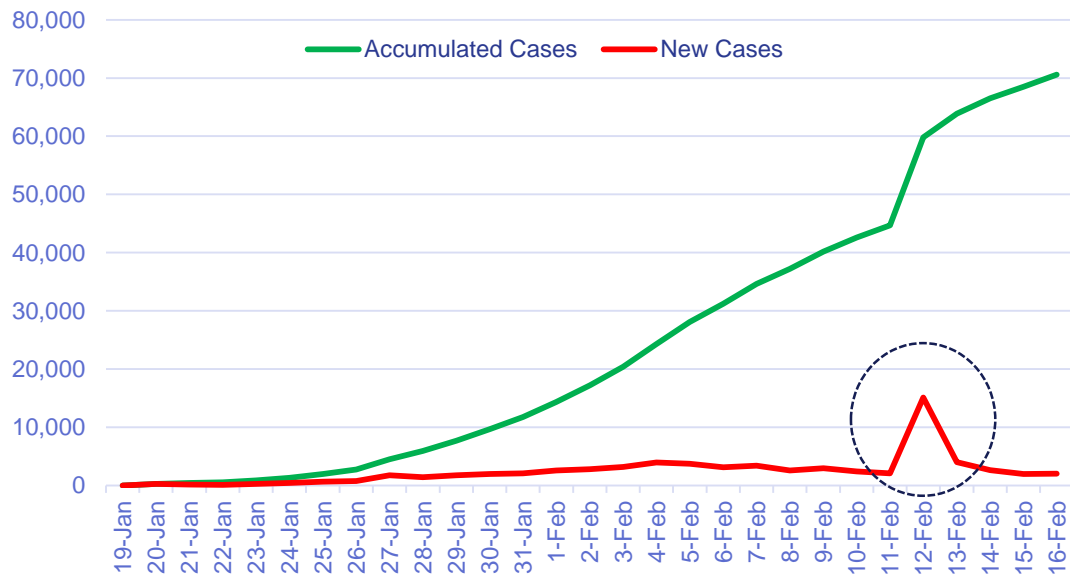


# Daily Confirmed Virus Cases

City\_Virus\_20200204.xlsx - Excel

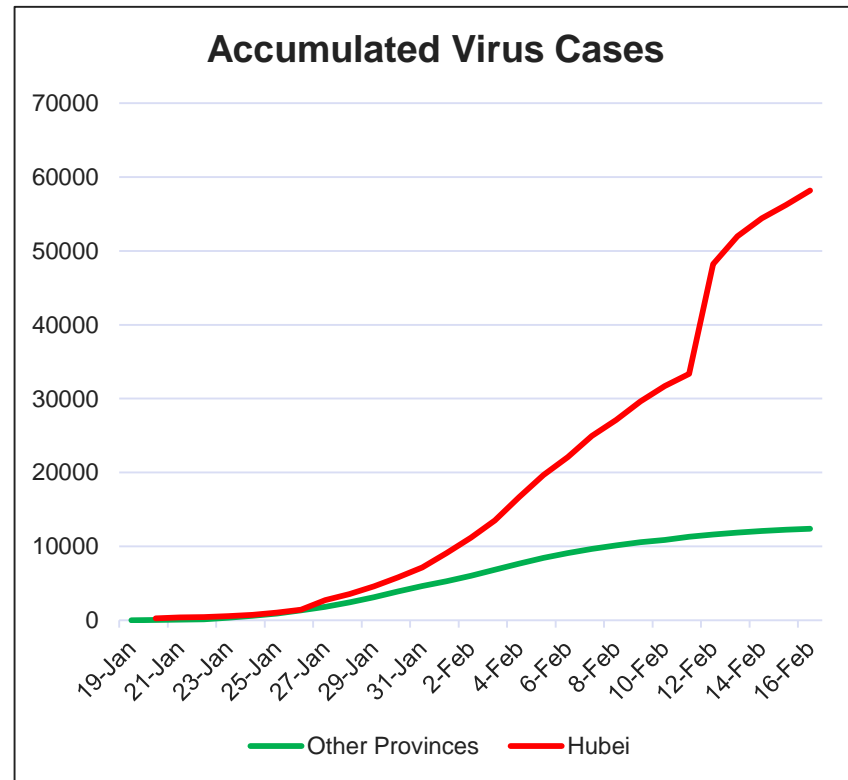
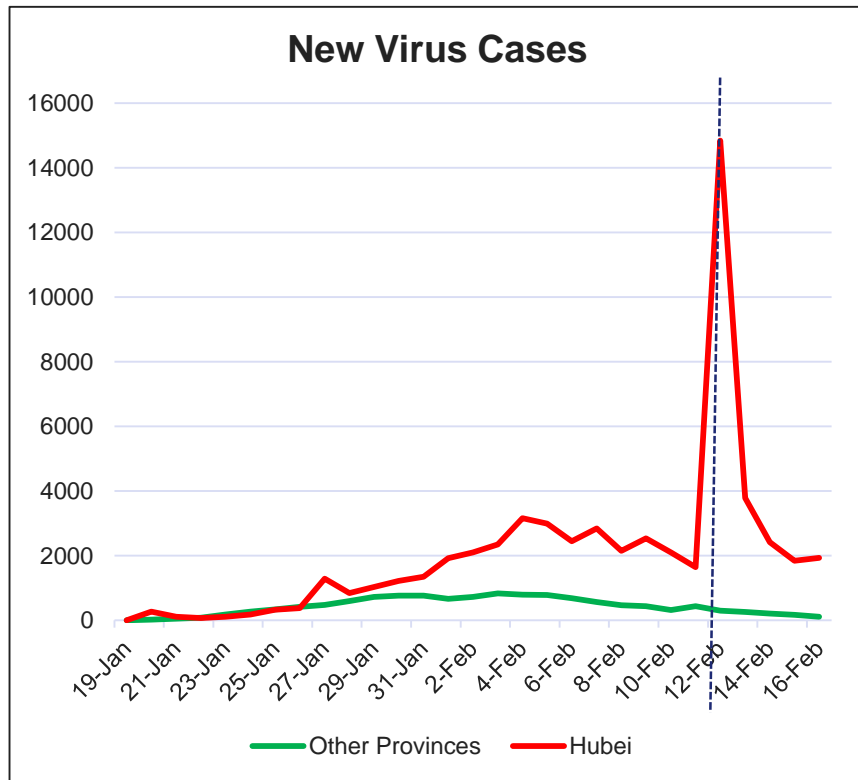
	E	F	G	AB	AC	AD	AE	AF
1	GbCity	City_CH	City_EN	Confirm_1	Confirm_2	Confirm_2	Confirm_2	Confirm_2
174	4201	武汉市	Wuhan	3215	4109	5142	6384	8351
175	4202	黄石市	Huangshi	209	252	334	405	509
176	4203	十堰市	Shiyan	177	212	256	291	318
177	4205	宜昌市	Yichang	276	353	392	452	496
178	4206	襄樊市	Xiangfan	347	441	548	632	735
179	4207	鄂州市	Ezhou	227	278	306	332	382
180	4208	荆门市	Jingmen	251	329	345	400	422
181	4209	孝感市	Xiaogan	628	749	918	1120	1462
182	4210	荆州市	Jingzhou	287	333	499	613	713
183	4211	黄冈市	Huanggang	726	1002	1246	1422	1645
184	4212	咸宁市	Xianning	206	246	296	348	384
185	4213	随州市	Suizhou	304	384	458	641	706
186	4228	恩施土家	Enshitujiazum	87	105	111	123	138
187	4290	省直辖行	Hubei (Direct					
188	4301	长沙市	Changsha	86	112	125	148	164
189	4302	株洲市	Zhuzhou	22	25	26	34	38
190	4303	湘潭市	Xiangtan	13	16	19	21	23
191	4304	衡阳市	Hengyang	26	30	32	35	39
192	4305	邵阳市	Shaoyang	41	49	51	55	68
193	4306	岳阳市	Yueyang	45	53	70	83	89
194	4307	常德市	Changde	39	42	45	50	56
195	4308	张家界市	Zhangjiajie					

## Daily Updates of New and Accumulated Confirmed Virus Cases in China



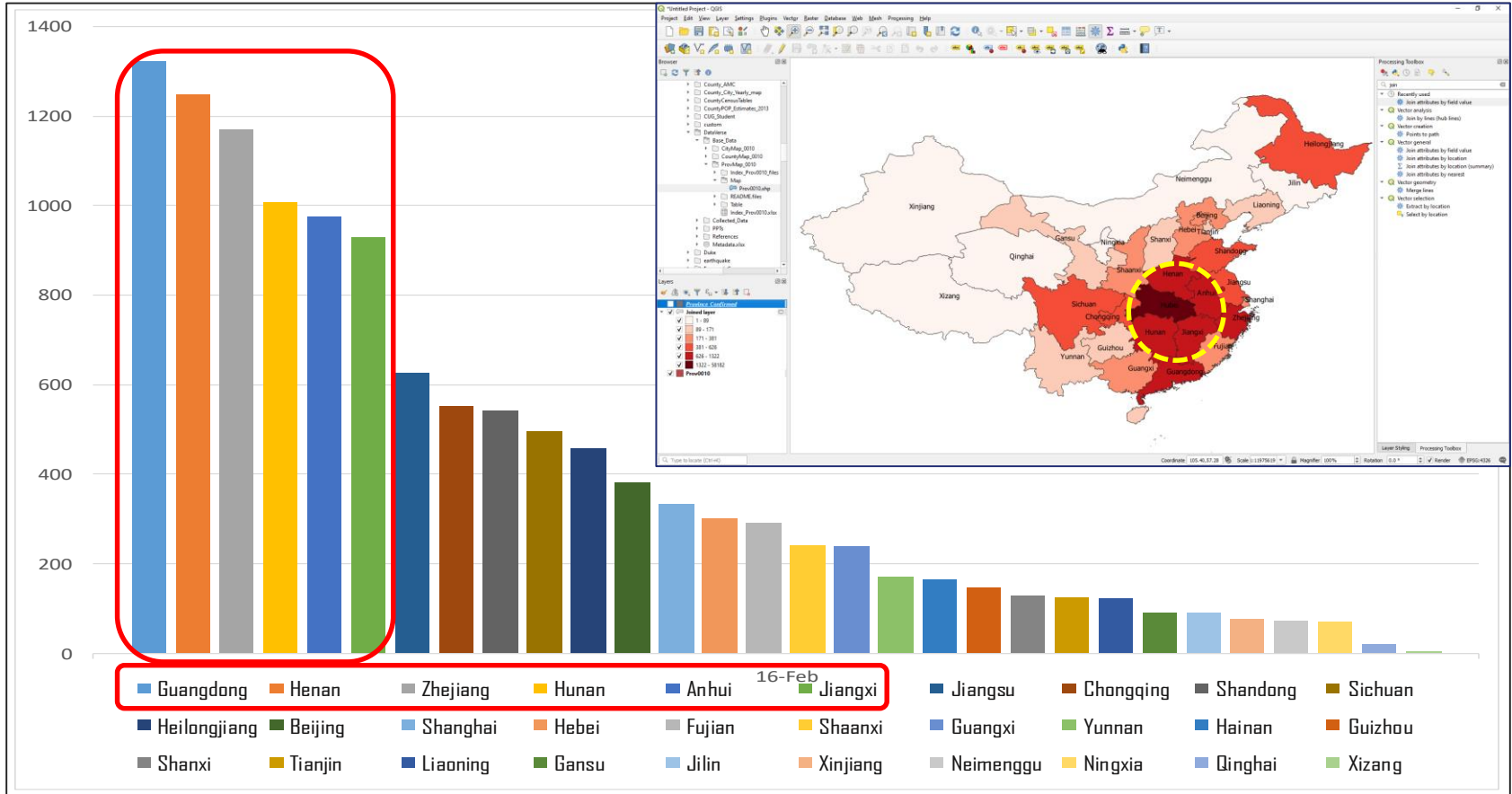
Data source: Ding Xiang Yuan, Real-time virus report ( <https://ncov.dxy.cn/ncovh5/view/pneumonia> )

# Read the Virus Data: Hubei & Other Provinces



Data source: Ding Xiang Yuan, Real-time virus report ( <https://ncov.dxy.cn/ncovh5/view/pneumonia> )

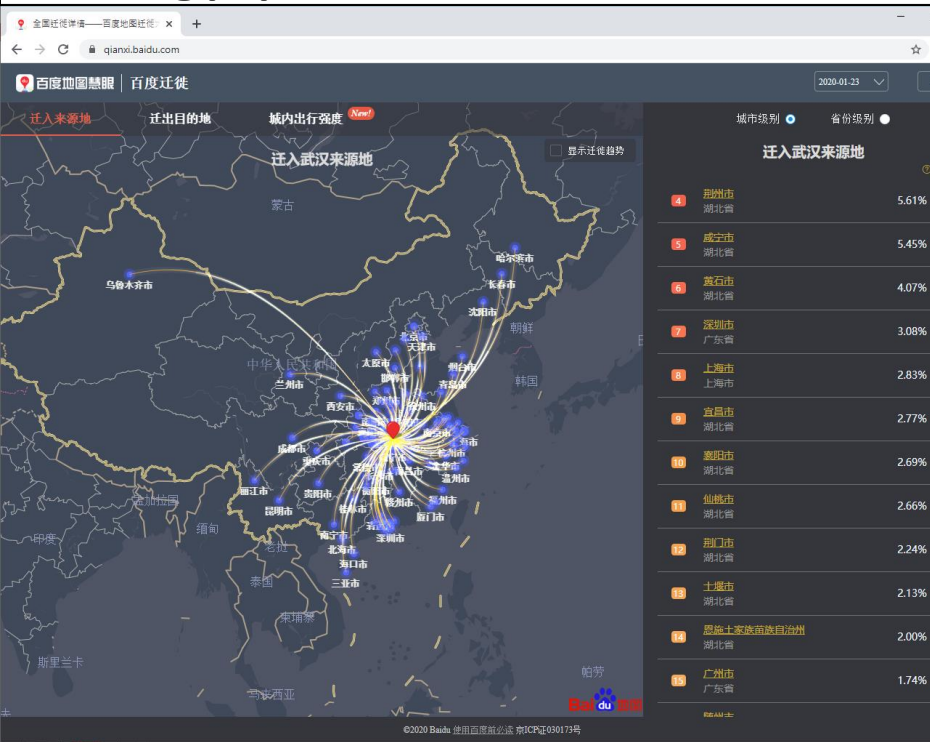
# Confirmed Virus Cases by Province As of Feb 16, 2020



# Daily Floating Population of Wuhan As of Jan 23, 2020

Data Source: <https://qianxi.baidu.com>

## Floating population from other cities to Wuhan



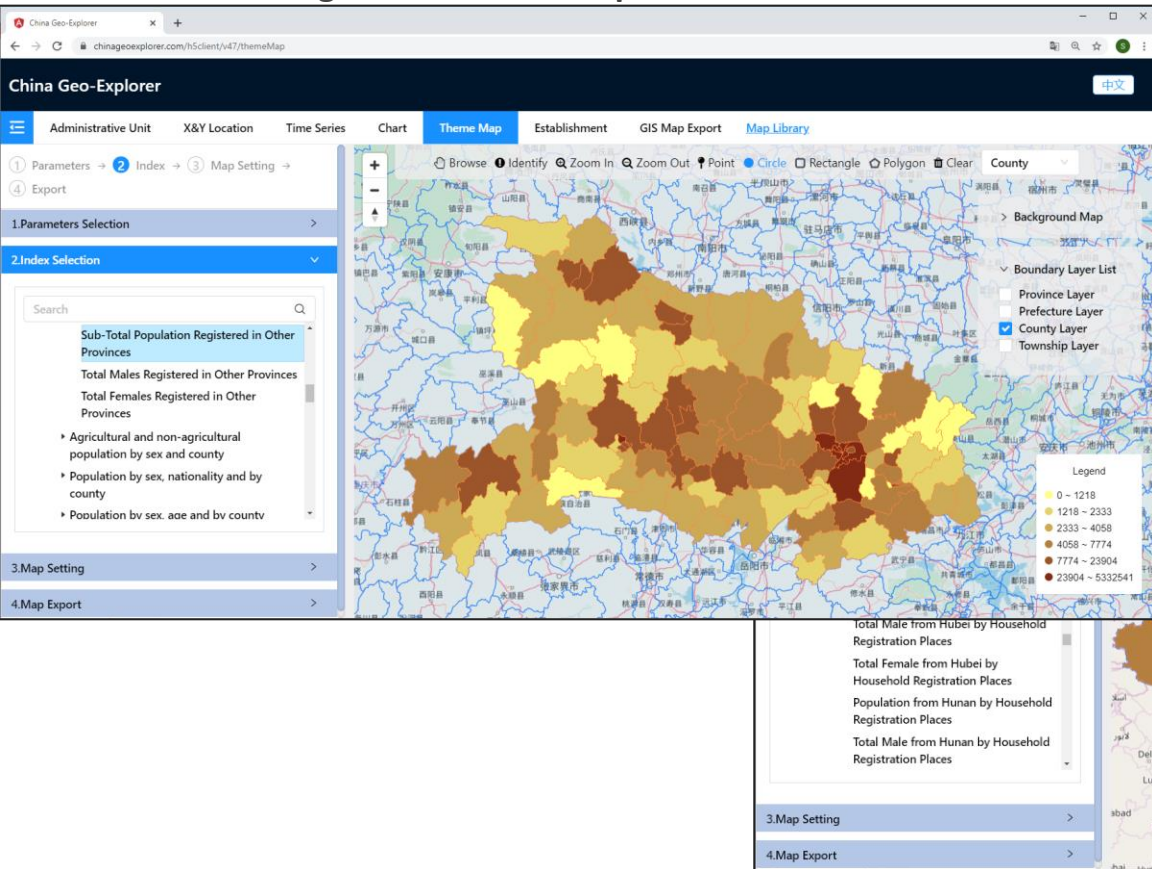
## Floating population from Wuhan to other cities



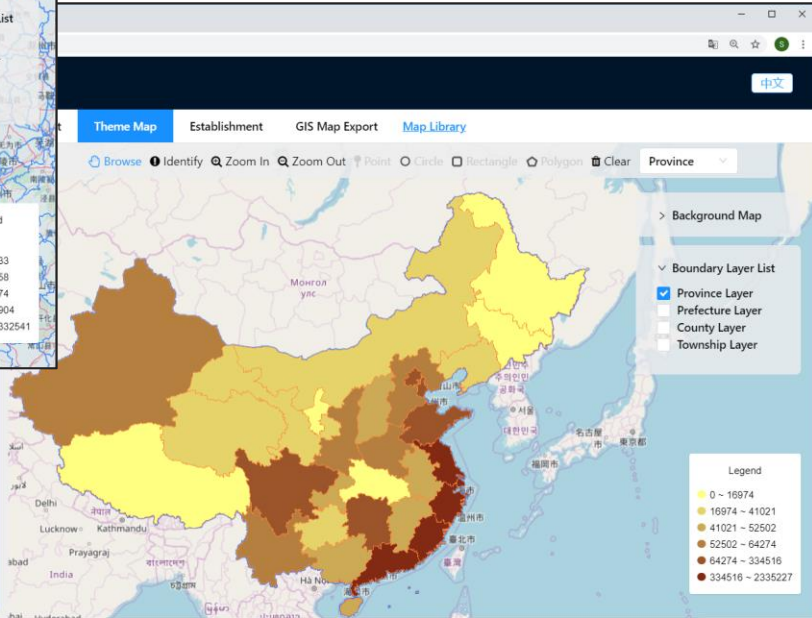
# Migration Data from 2010 Population Census (> 6 ms & < 5 ys)

Data Source: <https://china-data-online.com>

## Migration from other provinces to Hubei

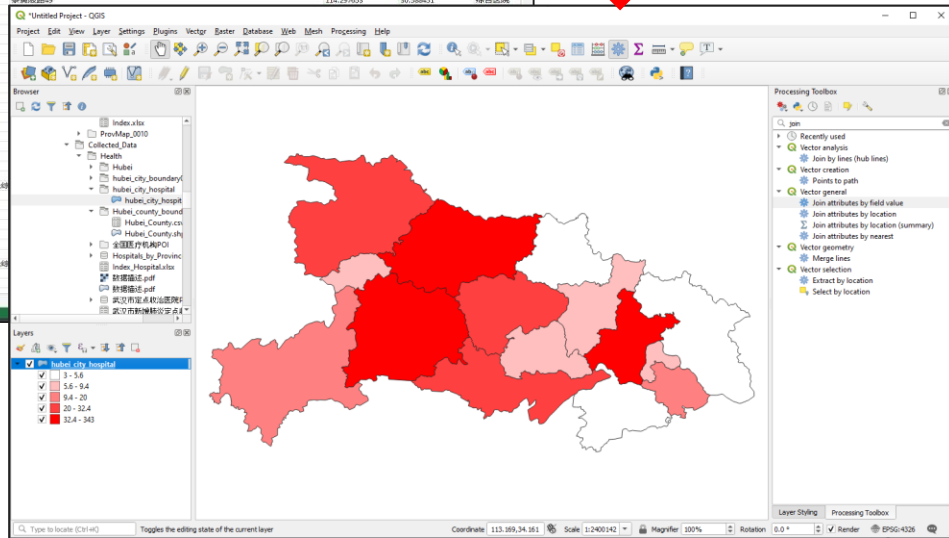
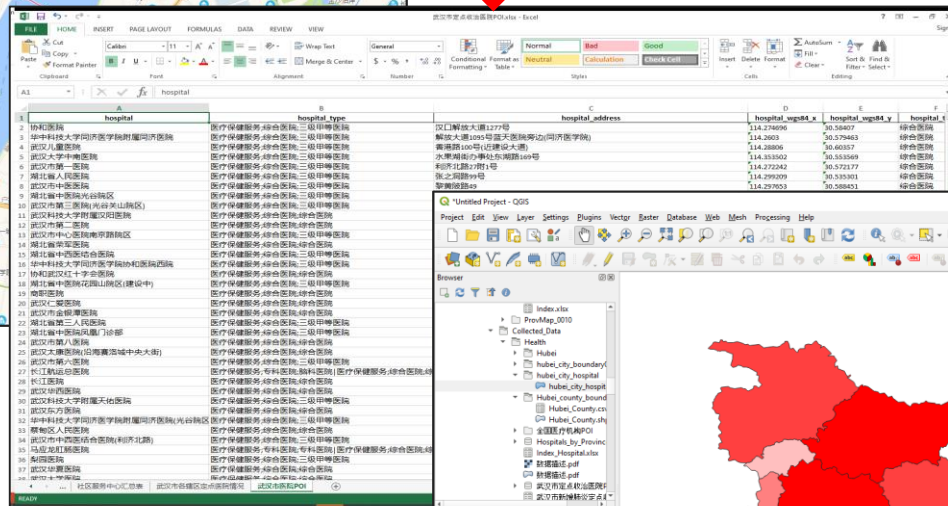
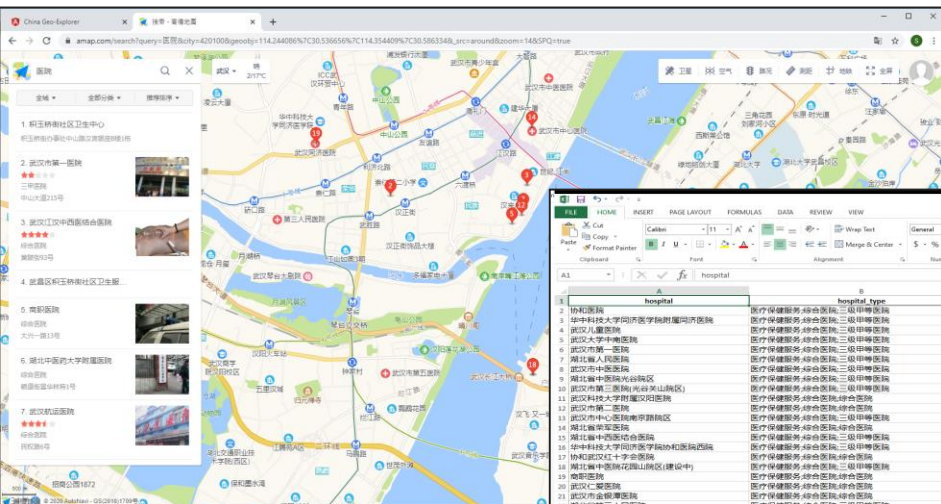


## Migration from Hubei to other provinces



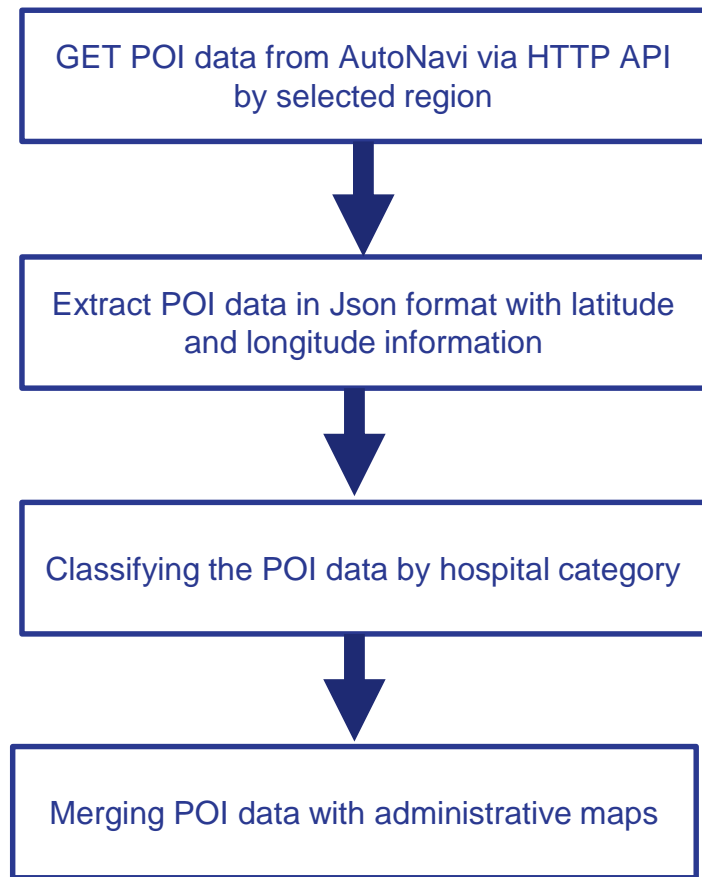
# Hospital Data of Wuhan from AutoNavi

Data Source: <https://lbs.amap.com>



# Hospital Data Retrieval from AutoNavi

Data Source: <https://lbs.amap.com>

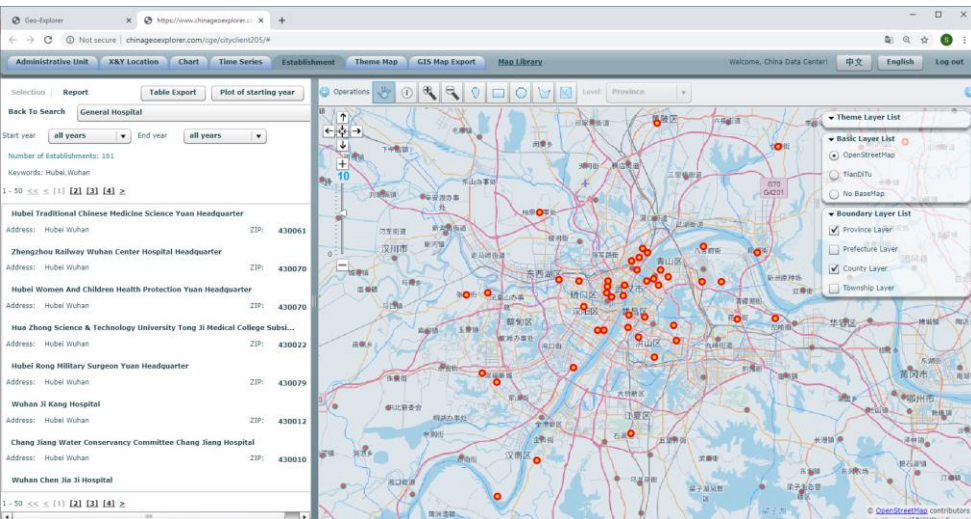


Name_CH	Name_EN
医院总数	Total Hospital Number
三级甲等医院	Tertiary Hospital
专科医院	Specialist Hospital
传染病医院	Infectious Disease Hospital
医疗保健服务场所	Healthcare Services
卫生院	Health Center
口腔医院	Dental Hospital
妇科医院	Women's Hospital
急救中心	Emergency Center
整形美容	Plastic Surgery
疾病预防	Prevent Disease Center
眼科医院	Eye Hospital
精神病医院	Psychiatric Hospital
综合医院	general Hospital
耳鼻喉医院	ENT Hospital
肿瘤医院	Cancer Hospital
胸科医院	Chest Hospital
脑科医院	Brain Hospital
诊所	Clinic
骨科医院	Orthopaedic Hospital

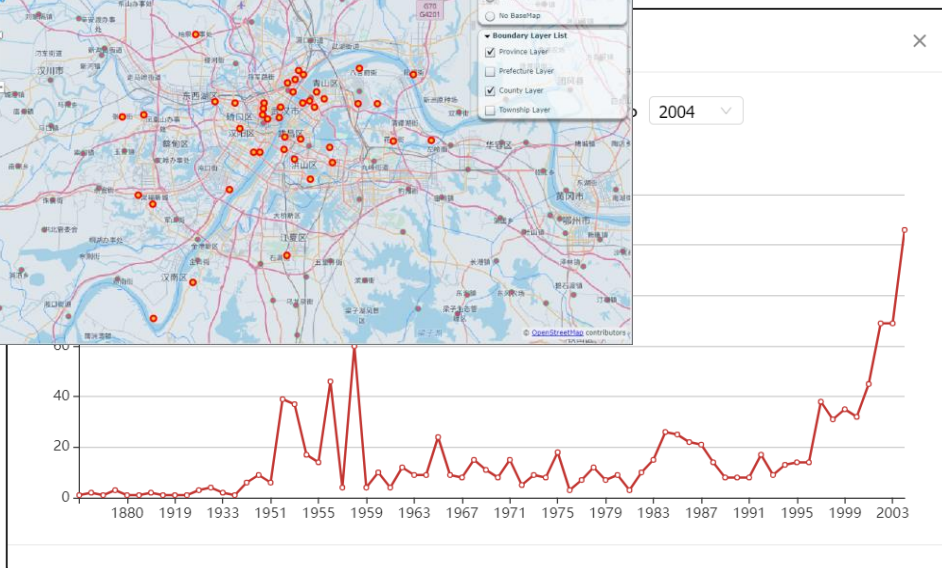
# Hospital Data from Economic Census (2004 & 2008)

Data Source: <https://china-data-online.com>

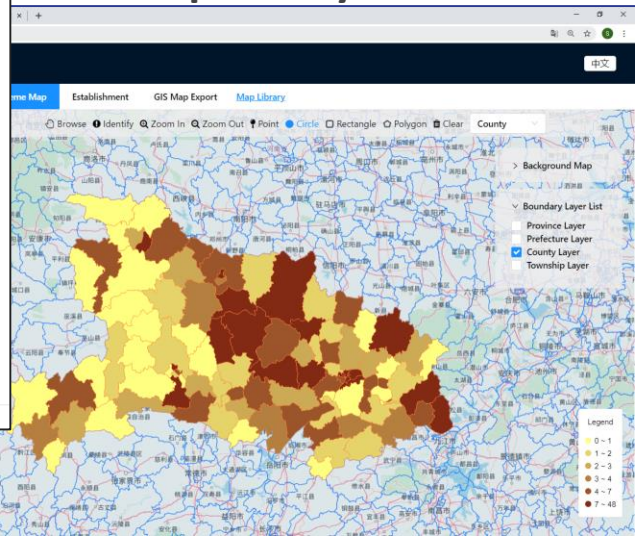
## Locations of general hospitals in Wuhan



## New hospitals by year



## Hospitals by Counties

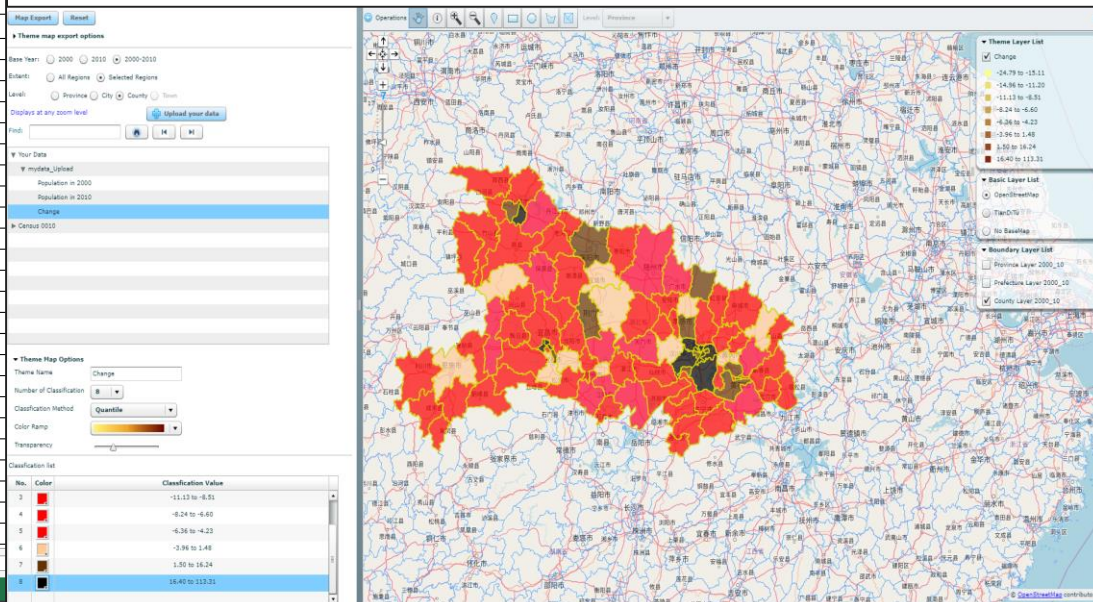


# Population Changes from 2000 to 2010

Data Source: <https://china-data-online.com>

A	B	C
4 Prov_EN	所在省的英文名称	English name of Province
5 GbCity	市级的四位码	Four digit code of City in number
6 City_CH	所在市的中文名称	Chinese name of City
7 City_EN	所在市的英文名称	English name of City
8 GbCounty	区县级的六位码	Six digit code of County in number
9 County_CH	所在区县的中文名称	Chinese name of County
10 County_EN	所在区县的英文名称	English name of County
11 A100008_00	2000年城镇人口数	Urban Population in 2000
12 A100008_10	2010年城镇人口数	
13 A100009_00	2000年乡村人口数	
14 A100009_10	2010年乡村人口数	
15 A101001_00	2000年总户数合计	
16 A101001_10	2010年总户数合计	
17 A101002_00	2000年属家庭户的户数	
18 A101002_10	2010年属家庭户的户数	
19 A101003_00	2000年属集体户的户数	
20 A101003_10	2010年属集体户的户数	
21 A101004_00	2000年总人口数合计	
22 A101004_10	2010年总人口数合计	
23 A101005_00	2000年男性人口数合计	
24 A101005_10	2010年男性人口数合计	
25 A101006_00	2000年女性人口数合计	
26 A101006_10	2010年女性人口数合计	
27 A103001_00	2000年户籍人口数	
28 A103001_10	2010年户籍人口数	
29 A104001_00	2000年户口登记地在外乡镇街道的总人口合计	
30 A104001_10	2010年户口登记地在外乡镇街道的总人口合计	
31 A104004_00	2000年户口登记地在本县(市、区)的人口小计	
32 A104004_10	2010年户口登记地在本县(市、区)的人口小计	
33 A104007_00	2000年户口登记地在本省其他县(市、区)的人口小计	
34 A104007_10	2010年户口登记地在本省其他县(市、区)的人口小计	
35 A104010_00	2000年户口登记地在省外的人口小计	
36 A104010_10	2010年户口登记地在省外的人口小计	
37 A105004_00	2000年农业户口人口小计	
38 A105004_10	2010年农业户口人口小计	
39 A105007_00	2000年非农业户口人口小计	
40 A105007_10	2010年非农业户口人口小计	

## Population Changes by County: 2000 - 2010



# Health Statistics on China Data Online

<http://china-data-online.com>

The screenshot shows the homepage of the China Data Online website. The header includes the logo for ALL CHINA DATA CENTER (ACMR) and the text 'China Data Online 中国数据在线'. Below the header is a navigation menu with links for Home, Data Products, Database Demo, Dictionary, Support, Contact, Q&A, Citations, My Account, and Logout. The main content area is divided into three red-bordered boxes:

- CHINA SPATIAL DATA**
  - China Geo-Explorer II
  - China Geo-Explorer I
  - China Map Library
- CHINA STATISTICS**
  - Monthly Statistics
  - National Statistics
  - Provincial Statistics
  - City Statistics
  - County Statistics
  - Monthly Industrial Data
  - Yearly Industrial Data
  - Statistics on Map
  - Statistical Datasheets
  - Statistical Charts
- CENSUS DATA**
  - Census Maps
  - All Census Data
  - Economic Census 2004
  - Industrial Census 1995
  - Census 1982
  - Census 1982 (10%)
  - Census 1990
  - Census 1995 (1%)
  - Province 2000
  - County 2000
  - Census 2005 (1%)
  - Census Data Search

At the bottom, there is a section for **FREE CHINA MAPS** with links to '2000 Population Census', 'Pop & Env (1990-1999)', 'Pop & Env (2000)', and 'Atlas of Industrial Census'. A footer link reads 'The Gross Imports and Exports in the Fourth Quarter of 2019 (1/14/2020)'.


## Health Statistics

- National Statistics
- Provincial Statistics
- City Statistics
- County Statistics
- Monthly Industrial Data
- Yearly Industrial Data
- Statistical Datasheets

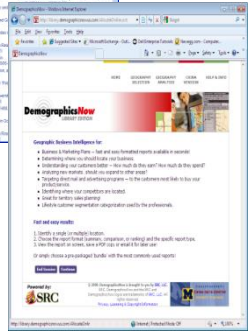
# Health Data on China Geo-Explorer

<https://china-data-online.com>


**Statistics**



**Census**



**GIS**



Data



**All China Marketing Research**

Welcome China Data Center, your IP is: 98.224.227.102, please click the links below to enter.

➔ **China Geo-Explorer II**

China Geo-Explorer (CGE) fully integrates different data sources from government statistics, population census and economics census of China at different levels (province, city, county, township and ZIP code) into a spatial system with more than 6,000 comparable variables for easy access.

[Full\\_Version](#) [HTML5\\_Beta\\_Version](#) [Free\\_Version](#) [User\\_Guide](#) [Citations](#)

➔ **US Geo-Explorer**

US Geo-Explorer (UGE) fully integrates the population census and business data of the U.S. from different years at different levels (state, metropolitan, county, CCD, place, tract and block) into a spatial system with more than 40,000 comparable variables for easy access.

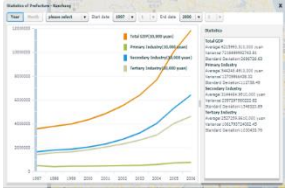
[User's\\_Guide](#)

\*Both systems are compatible with IE 8+, FireFox or Chrome. Flash player 9 or higher version is required.


All China Market Research (ACMR)

Output

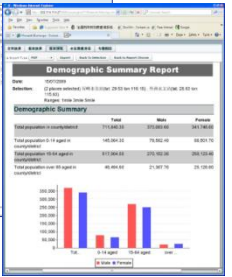
**Charts**



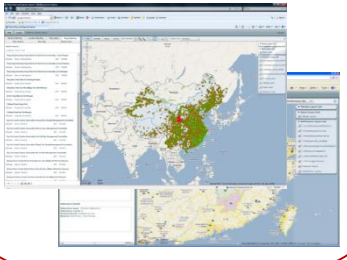
**Tables**



**Reports**



**Maps**





# Yearly Statistical Datasheets

Data Source: <https://china-data-online.com>

Welcome UM Asia Library , your IP is: 141.213.168.252 Today is: 2016-11-20 EST USA

**ALL CHINA DATA CENTER** China Data Online 中国数据在线

Home | CDC | Database Demo | Offline Products | Dictionary | Support | Contact | Online Question And Answer | My Account | Logout

Region:  CHINA  BEIJING  TIANJIN  HEBEI  SHANXI  INNER MONGOLIA  LIAONING  JILIN  HEILONGJIANG  SHANGHAI  JIANGSU  ZHEJIANG  ANHUI  FUJIAN  JIANGXI  SHANDONG  HENAN  HUBEI  HUNAN  GUANGDONG  GUANGXI  HAINAN  CHONGQING  SICHUAN  GUIZHOU  YUNNAN  TIBET  SHAANXI  GANSU

Year: 2014 To 2015 hospital Search In Results Search

Datasheet Name	Region	Subject	Year
Hospital Patients (2013)	HEBEI	PUBLIC HEALTH AND SOCIAL ...	2014
Operation of Hospitals, Health Care Centers	TIANJIN	PUBLIC HEALTH AND SOCIAL ...	2014
20-4 BASIC STATISTICS ON HOSPITALS (2013)	BEIJING	PUBLIC HEALTH AND SOCIAL ...	2014
20-5 WORKS OF HOSPITALS (2013)	BEIJING	PUBLIC HEALTH AND SOCIAL ...	2014
18-8 Medical Services of Hospitals in Selected Years	FUJIAN	PUBLIC HEALTH AND SOCIAL ...	2014
18-9 Medical Services of Hospitals, Institutes of Health and Health...	FUJIAN	PUBLIC HEALTH AND SOCIAL ...	2014
Number of Hospital Patients & Admissions (2013)	GUANGXI	PUBLIC HEALTH AND SOCIAL ...	2014
Basic Statistics on Hospitals and Health Centers at County Level ...	GUIZHOU	PUBLIC HEALTH AND SOCIAL ...	2014
Services in Hospitals by Region	GUIZHOU	MAIN INDICATORS OF CITIES...	2014
Percentage of 10 Main Diseases of Inpatients in City Hospitals of ...	HENAN	PUBLIC HEALTH AND SOCIAL ...	2014
Hospital Beds and Medical Technical Personnel by City and County ...	XINJIANG	PUBLIC HEALTH AND SOCIAL ...	2014
Hospital Beds and Medical Technical Personnel by City and County ...	XINJIANG	PUBLIC HEALTH AND SOCIAL ...	2014
The per capita medical expenses of patients hospitalized in the m...	XINJIANG	PUBLIC HEALTH AND SOCIAL ...	2014
The per capita medical expenses of patients hospitalized in the m...	XINJIANG	PUBLIC HEALTH AND SOCIAL ...	2014
Number of Doctors and Hospital Beds Per 10 000 Population by Pref...	XINJIANG	MAIN INDICATORS OF CITIES...	2014
Statistics on Visits and Inpatients in Hospitals, Health Stations...	CHONGQING	PUBLIC HEALTH AND SOCIAL ...	2014
20-9 Number of Hospital Patients (2013)	ANHUI	PUBLIC HEALTH AND SOCIAL ...	2014
20-11 Utilization of Hospital Beds at and Above County Level (201...	ANHUI	PUBLIC HEALTH AND SOCIAL ...	2014
20-12 Hospital Beds Usage in Main Year	ANHUI	PUBLIC HEALTH AND SOCIAL ...	2014
20-13 Hospital Beds Usage by Region (2013)	ANHUI	PUBLIC HEALTH AND SOCIAL ...	2014
20-14 Hospital Beds Usage of Beds of Township Hospitals in Main Y...	ANHUI	PUBLIC HEALTH AND SOCIAL ...	2014
20-15 Hospital Beds Usage of Beds of Township Hospitals by Region...	ANHUI	PUBLIC HEALTH AND SOCIAL ...	2014
Situation of Outpatient and Hospitalization Services of Health Ins...	GANSU	PUBLIC HEALTH AND SOCIAL ...	2014
Basic Statistics of Hospitals and Health Centers (2013)	GANSU	PUBLIC HEALTH AND SOCIAL ...	2014
Out-patient Clinics in Hospitals in Medical Institutions(2013)	JIANGXI	PUBLIC HEALTH AND SOCIAL ...	2014

Copyright © 2005 - 2011 All China Marketing Research Co., Ltd. 京ICP证050875号 Product Introduction Contact Version 4.01

# Manufacture of Medical Machinery

Data Source: <https://china-data-online.com>

医疗诊断、监护及治疗设备制造	Medical diagnosis, care and treatment equipment manufacturing
口腔科用设备及器具制造	Dental equipment and apparatus used manufacturing
实验室及医用消毒设备和器具的制	Laboratory and medical equipment and apparatus disinfection system
医疗、外科及兽医用器械制造	Medical, surgical and veterinary equipment manufacturing
机械治疗及病房护理设备制造	Treatment and Nursing mechanical equipment manufacturing
假肢、人工器官及植（介）入器械	artificial organs and plantations (referred) to enter Devices
其他医疗设备及其器械制造	Other medical equipment and device manufacturing

	Baoshan	Changning	Changshu	Chongming	Fengxian	Hongkou	Huangpu	Jiading	Jingan	Kunshan	Luwan	Mixing	Nanhui
2008 Daily and medical rubber products manufacturing	4	2	5	0	3	2	0	14	0	4	0	4	4
2008 Medical diagnosis, care and treatment equipment manufacturing	8	0	2	4	9	5	0	12	2	4	2	8	10
2008 Dental equipment and apparatus used manufacturing	1	0	1	2	3	1	0	1	0	2	0	4	4
2008 Laboratory and medical equipment and apparatus disinfection system	6	1	1	1	3	1	0	2	0	2	0	3	2
2008 Medical, surgical and veterinary equipment manufacturing	10	2	6	0	6	3	2	13	0	8	0	8	12
2008 Treatment and Nursing mechanical equipment manufacturing	0	2	0	0	3	2	0	1	0	1	0	3	1
2008 artificial organs and plantations (referred) to enter Devices	3	5	1	0	2	4	0	0	7	0	1	8	2
2008 Other medical equipment and device manufacturing	11	1	9	5	11	4	0	14	0	6	1	16	7

# Manufacture of Medicines

Data Source: <https://china-data-online.com>

化学药品原药制造	The original drug manufacturing chemicals
化学药品制剂制造	Chemical agent production
中药饮片加工	Pieces of Traditional Chinese medicine processing
中成药制造	Traditional Chinese Prepared Medicines
兽用药品制造	Veterinary medicine manufacturing
生物、生化制品的制造	Biological, chemical and biological products manufacturing
卫生材料及医药用品制造	Sanitation Materials and Medical Articles

	Baoshan	Changning	Changshu	Chongming	Fengxian	Hongkou	Huangpu	Jiading	Jingan	Kunshan
2008 The original drug manufacturing chemicals	5	0	12	0	5	1	1	13	0	20
2008 Chemical agent production	5	2	2	1	11	0	0	14	0	9
2008 Pieces of Traditional Chinese medicine processing	1	0	0	1	1	1	1	4	0	0
2008 Traditional Chinese Prepared Medicines	2	0	2	2	3	0	1	5	1	0
2008 Veterinary medicine manufacturing	4	0	1	0	3	0	0	4	0	0
2008 Biological, chemical and biological products manufacturing	9	1	2	0	17	1	1	3	0	9
2008 Sanitation Materials and Medical Articles	15	2	6	4	12	0	0	5	1	13

# Health Care

Data Source: <https://china-data-online.com>

综合医院  
中医医院  
中西医结合医院  
民族医院  
专科医院  
疗养院  
卫生院及社区医疗活动  
门诊部医疗活动  
计划生育技术服务活动  
妇幼保健活动  
专科疾病防治活动  
疾病预防控制及防疫活动  
其他卫生活动  
社会保障业  
干部休养所  
收养收容服务  
不提供住宿的社会福利

General Hospital  
Traditional Chinese medicine hospital  
Combined Traditional Chinese and Western Medicine Hospital  
National Minority Hospital  
Specialist Hospital  
Nursing Homes  
Health Centers and Community Medical Treatment Activities  
Clinics Medical Treatment Activities  
Family planning service  
Maternal and child health service  
Specialist service for Disease Control and Prevention  
Vaccination for disease prevention and control service  
Other health services  
Social Security  
Cadres sanatorium  
Adoption of asylum services  
No Lodging Social Welfare

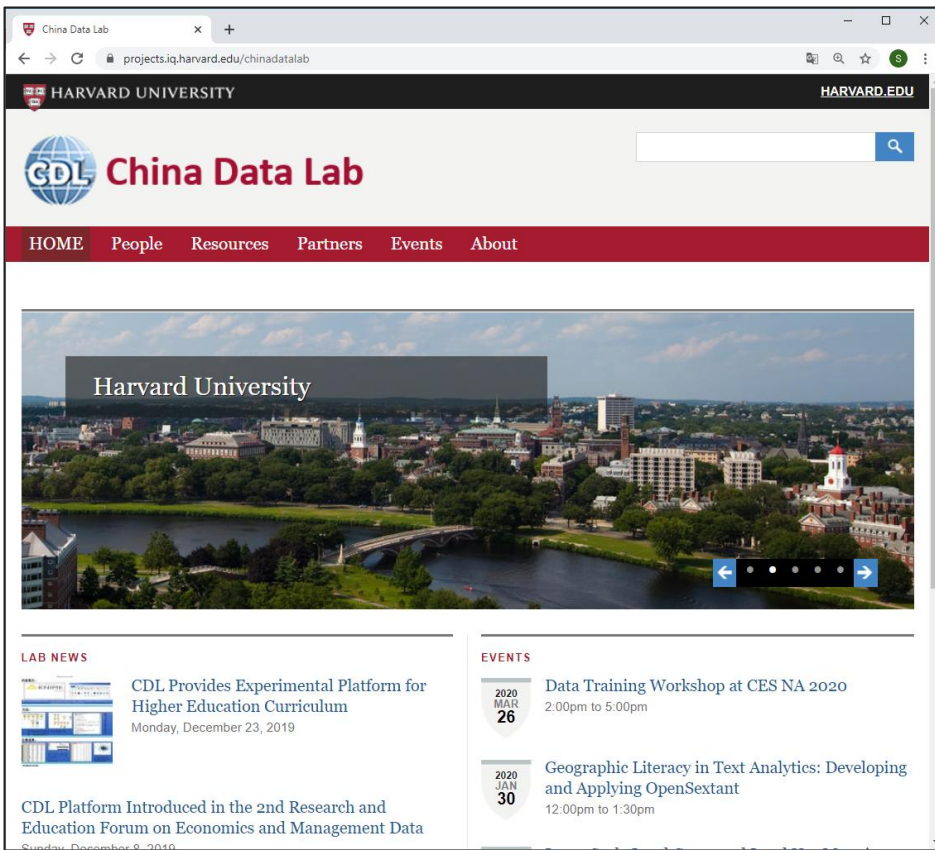
Date:	14/10/2016		
Selection:	(24 selected) Baoshan, Changning, Changshu, Chongming, Fengxian, Hongkou, Huangpu, Jiading		
	Baoshan	Changning	Changshu
2008 General Hospital	14	12	8
2008 Traditional Chinese medicine hospital	1	2	1
2008 Combined Traditional Chinese and Western Medicine Hospital	0	1	0
2008 National Minority Hospital	0	0	0
2008 Specialist Hospital	5	17	11
2008 Nursing Homes	2	5	1
2008 Health Centers and Community Medical Treatment Activities	18	7	27
2008 Clinics Medical Treatment Activities	22	31	22
2008 Family planning service	2	3	8
2008 Maternal and child health service	2	2	1
2008 Specialist service for Disease Control and Prevention	3	3	0
2008 Vaccination for disease prevention and control service	1	6	4
2008 Other health services	6	5	10

# Output: The List of Some Data Products From this Project

- Daily reports on virus by city
- Daily reports on virus by province
- Daily reports on virus by country
- City GIS Maps with Daily Virus Data
- Province GIS Maps with Daily Virus Data
- Country GIS maps with daily virus data
- Workflow data analysis with virus data
- Workflow data analysis with social media data
- Matrix tables with daily city to city floating population
- Matrix tables with daily province to province floating population
- County GIS maps with hospital data of China
- City GIS maps with hospital data of China
- Province GIS maps with hospital data of China
- Policies, regulations and changes in leadership

# Deployment I: Resource Center for COVID-19

<http://chinadatalab.net>



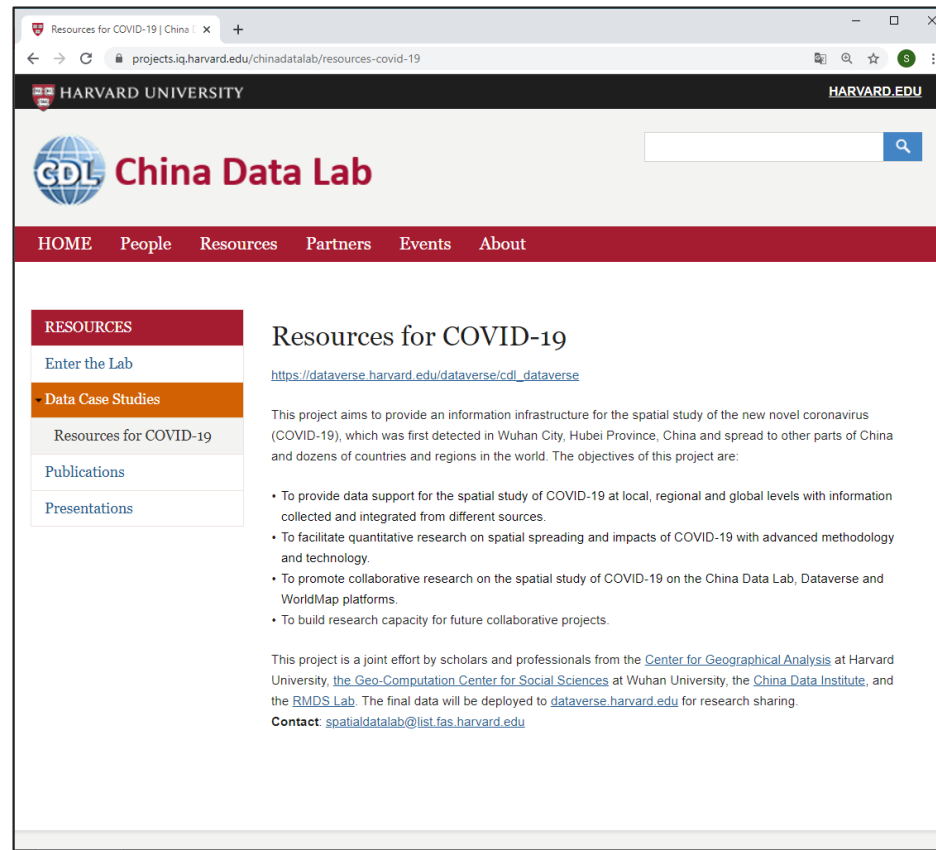
The screenshot shows the homepage of the China Data Lab. The browser address bar displays `projects.iq.harvard.edu/chinadatalab`. The page features the Harvard University logo and the CDL China Data Lab branding. A navigation menu includes links for HOME, People, Resources, Partners, Events, and About. A large banner image of Harvard University is displayed with the text "Harvard University". Below the banner, there are sections for "LAB NEWS" and "EVENTS".

**LAB NEWS**

- CDL Provides Experimental Platform for Higher Education Curriculum**  
Monday, December 23, 2019
- CDL Platform Introduced in the 2nd Research and Education Forum on Economics and Management Data**  
Sunday, December 8, 2019

**EVENTS**

- 2020 MAR 26** Data Training Workshop at CES NA 2020  
2:00pm to 5:00pm
- 2020 JAN 30** Geographic Literacy in Text Analytics: Developing and Applying OpenSextant  
12:00pm to 1:30pm



The screenshot shows the "Resources for COVID-19" page. The browser address bar displays `projects.iq.harvard.edu/chinadatalab/resources-covid-19`. The page features the Harvard University logo and the CDL China Data Lab branding. A navigation menu includes links for HOME, People, Resources, Partners, Events, and About. The "RESOURCES" section is active, showing a list of resource categories: Enter the Lab, Data Case Studies, Resources for COVID-19, Publications, and Presentations. The "Resources for COVID-19" category is selected, displaying a list of resources and a contact information section.

**RESOURCES**

- Enter the Lab
- Data Case Studies
- Resources for COVID-19
- Publications
- Presentations

**Resources for COVID-19**

[https://dataverse.harvard.edu/dataverse/cdl\\_dataverse](https://dataverse.harvard.edu/dataverse/cdl_dataverse)

This project aims to provide an information infrastructure for the spatial study of the new novel coronavirus (COVID-19), which was first detected in Wuhan City, Hubei Province, China and spread to other parts of China and dozens of countries and regions in the world. The objectives of this project are:

- To provide data support for the spatial study of COVID-19 at local, regional and global levels with information collected and integrated from different sources.
- To facilitate quantitative research on spatial spreading and impacts of COVID-19 with advanced methodology and technology.
- To promote collaborative research on the spatial study of COVID-19 on the China Data Lab, Dataverse and WorldMap platforms.
- To build research capacity for future collaborative projects.

This project is a joint effort by scholars and professionals from the [Center for Geographical Analysis](#) at Harvard University, the [Geo-Computation Center for Social Sciences](#) at Wuhan University, the [China Data Institute](#), and the [RMDS Lab](#). The final data will be deployed to [dataverse.harvard.edu](https://dataverse.harvard.edu) for research sharing.

**Contact:** [spatialdatalab@list.fas.harvard.edu](mailto:spatialdatalab@list.fas.harvard.edu)

# Deployment II: Data Archives on dataverse.harvard.edu

<http://covid-19.chinadatalab.net>



Open source research data repository software



Researchers

Enjoy full control over your data. Receive *web visibility*, *academic credit*, and *increased citation counts*. A personal dataverse is easy to set up, allows you to display your data on your personal website, can be branded uniquely as your research program, makes your data more discoverable to the research community, and satisfies data management plans. *Want to set up your personal dataverse?*



Journals

Seamlessly manage the submission, review, and publication of data associated with published articles. Establish an *unbreakable link* between *articles in your journal* and *associated data*. Participate in the open data movement by using Dataverse as part of your journal data policy or list of repository recommendations. *Want to find out more about journal dataverses?*



Institutions

Establish a research data management solution for your community. Federate with a growing list of Dataverse repositories worldwide for increased discoverability of your community's data. Participate in the drive to set norms for sharing, preserving, citing, exploring, and analyzing research data. *Want to install a Dataverse repository?*



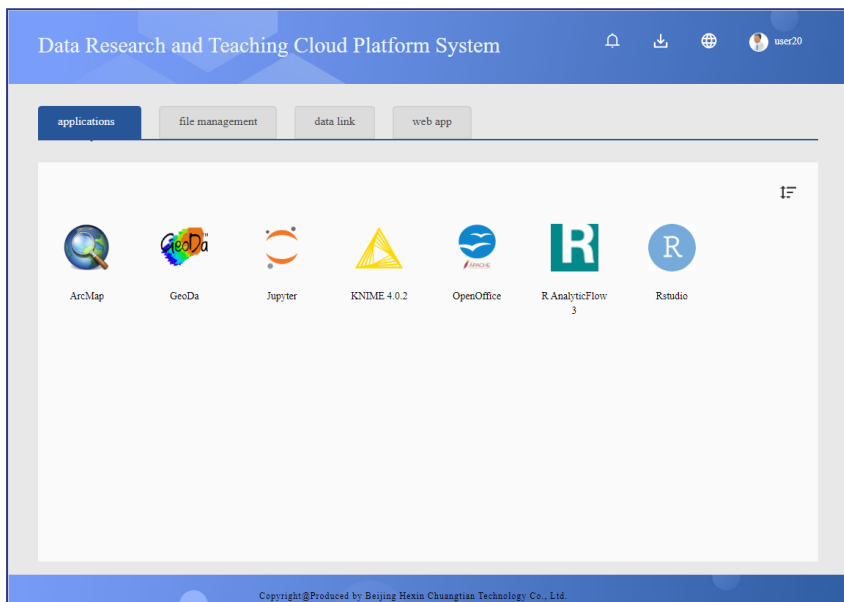
Developers

Participate in a vibrant and growing community that is helping to drive the norms for sharing, preserving, citing, exploring, and analyzing research data. Contribute code extensions, documentation, testing, and/or standards. *Integrate research analysis, visualization and exploration tools*, or other research and data archival systems with Dataverse. *Want to contribute?*

A screenshot of a web browser displaying the Dataverse website. The browser's address bar shows "dataverse.harvard.edu/dataverse/2019ncov". The page header includes the Harvard Dataverse logo and navigation links like "Add Data", "Search", "About", "User Guide", "Support", "Sign Up", and "Log In". The main heading is "Resources for COVID-19 (China Data Lab)". Below this is a navigation bar with four categories: "Data", "Development Code", "News Report", and "Research Papers". A search bar is present with a "Find" button and a link to "Advanced Search". The main content area shows a list of dataverses under the heading "1 to 6 of 6 Results". The list includes: "Data (China Data Lab) 2020-2-11", "Research Papers (China Data Lab) 2020-2-11", "Workflows (China Data Lab) 2020-2-11", "Web Sites (China Data Lab) 2020-2-11", and "News Report (China Data Lab) 2020-2-11". Each item has a small Dataverse icon to its right. A "Feedback" button is visible at the bottom right of the list.

# Deployment III: Executable Workflows on CDL Cloud

## CDL in Harvard



<http://harvard.chinadatalab.org>

## CDL in China

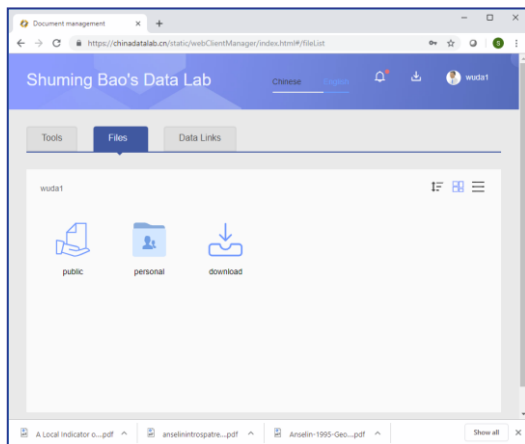


<http://chinadatalab.org>

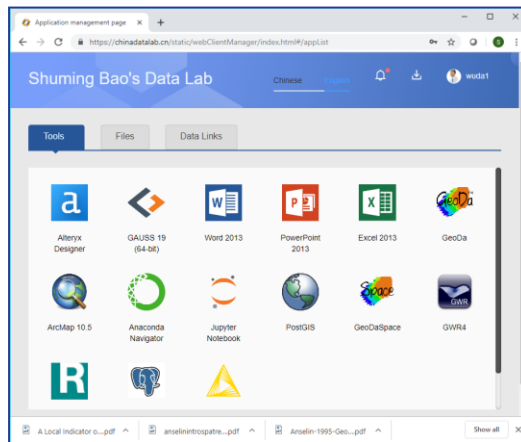
# CDL Cloud Platform Features

- ❑ Data available only on the cloud (users can upload own data)
- ❑ Tools available on the cloud
- ❑ All computation are on the cloud (the results can be downloadable)
- ❑ No maintenance required for end users

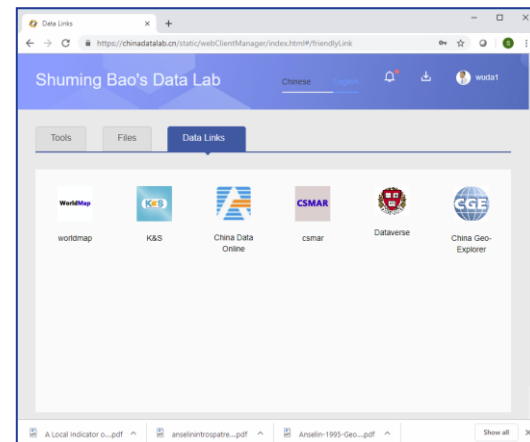
## Personal & Shared Data



## Tools



## External Data Links



# Tools on CDL Cloud

## Office tools



Word 2013



PowerPoint 2013



Excel 2013



Photoshop



Notepad++

## Interactive tools



GWR4



GeoDa



ArcMap 10.5



ENVI 5.3 + IDL 8.5



grass

## Programming tools



Rstudio



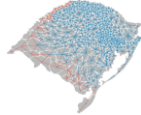
Anaconda Navigator



GAUSS 19

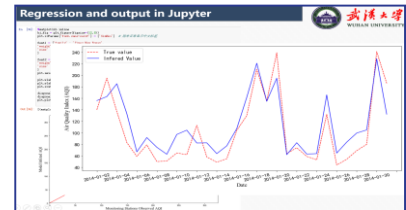
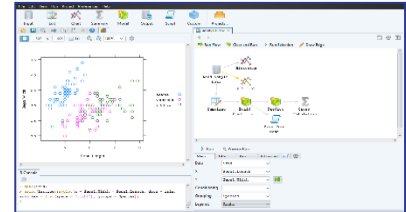
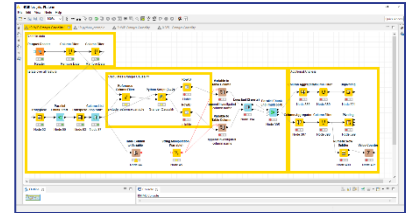


Spyder



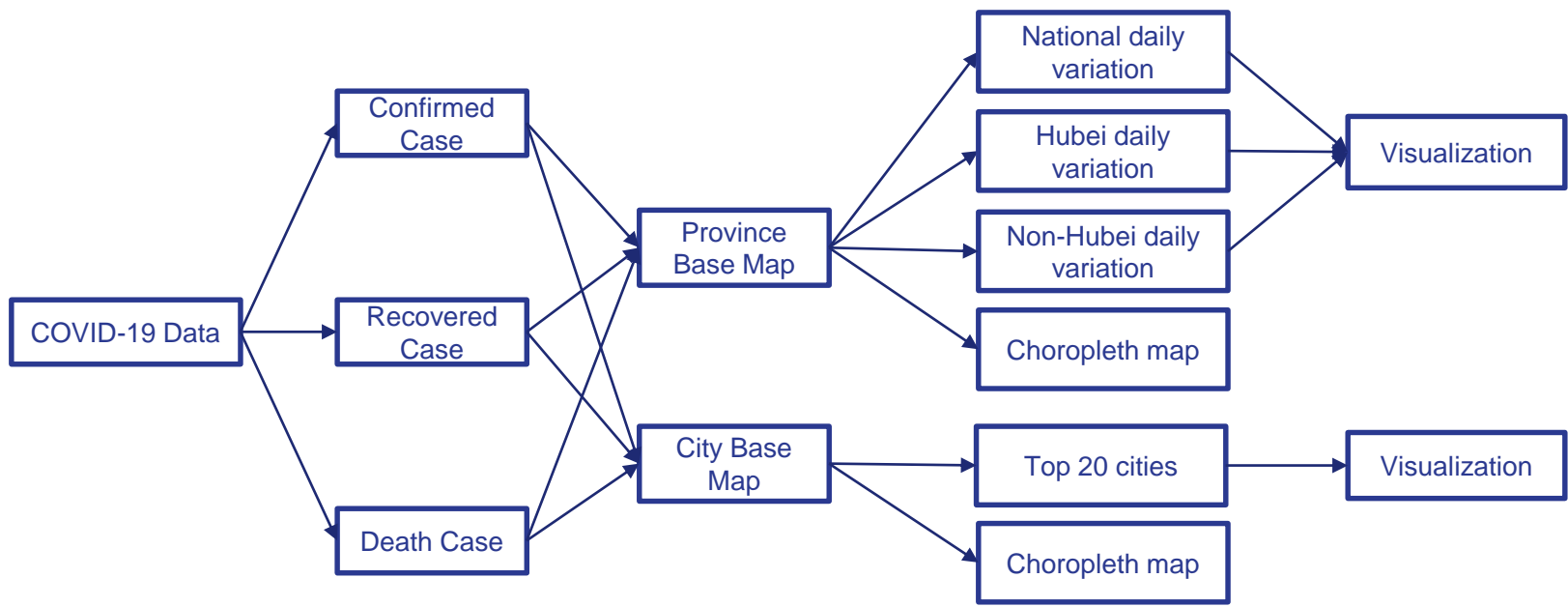
PySAL

## Workflow tools





# Flowcharts for the Integration, Analysis and Visualization of Virus Data



# Executable Workflow Data Analysis from DataVerse

Tool source: <https://www.knime.com/>



## Workflows for download on dataverse.harvard.edu

## Executable workflows on harvard.chinadatalab.org

**HARVARD Dataverse**  
Add Data Search About User Guide Support China Data Lab

Harvard Dataverse > China Data Lab Dataverse > Resources for COVID-19 > Workflows > COVID19 Stats Analysis

**COVID19 Stats Analysis**  
Draft Unpublished

Tao Hu, 2020, "COVID19 Stats Analysis", <https://doi.org/10.7910/DVNI/FWOPW2>, Harvard Dataverse, DRAFT VERSION

**Dataset Metrics**  
0 Downloads

**Description**  
This dataset saves workflows related to the COVID-19 statistics analysis. (2020-02-20)

**Subject**  
Medicine, Health and Life Sciences, Computer and Information Science, Social Sciences

**Keyword**  
workflow, statistics, map

Files Metadata Terms Versions

Search this dataset... Find

Filter by File Type: All Access: All

1 to 3 of 3 Files

- covid19\_stats\_confirmed.knwf  
 Unknown - 98.8 KB - 2020-2-20 - 0 Downloads  
 MDS: 301780709c5906c212ec3d66c07522c0  
 Copy and paste the following link to the executable workflow on your veappc://88D2AA413AED9040FAE30B4D4FBE0916B55E161735A7
- covid19\_stats\_death.knwf  
 Unknown - 99.7 KB - 2020-2-20 - 0 Downloads  
 MDS: 4b55d57e493618952ae1d4e0c26885  
 This workflow demonstrates the analysis of death cases in province veappc://88D2AA413AED9040FAE30B4D4FBE0916B55E161735A7
- covid19\_stats\_recovered.knwf  
 Unknown - 98.2 KB - 2020-2-20 - 0 Downloads  
 MDS: cfb5e9f1f29c99a0a01f7511744098  
 This workflow demonstrates the analysis of recovered cases in provi veappc://88D2AA413AED9040FAE30B4D4FBE0916B55E161735A7

Feedback

**Cases analysis in province level**

Row Filter (nation) → GroupBy (group by province) → Row Filter (hubei) → Row Filter (hubei) → Constant Value Column (add non-hubei) → GroupBy (igroup by province) → Concatenate (Node 9) → Transpose (combine) → Column Rename (copy column name) → Row Filter (delete 1st row) → RowID (delete 1st row) → String To Number (delete 1st row) → Line Plot (Node 22) → Image Writer (Port) (Export map)

**Choropleth map visualization in province level**

Province Map → Shapefile Polygon Reader → Column Filter (combination) → Joiner (category) → Rule Engine (set color) → Color Manager (group) → OSM Map to Image → OSM Viewer → Image Writer (Port) (Export map)

**Cases analysis in city level**

Row Filter (exclude nation) → GroupBy (group by province) → Sorter (Node 24) → Row Filter (Top 20 cities) → Transpose (transpose) → Insert Column Header (Node 156) → Row Filter (delete 1st row) → RowID (delete 1st row) → Column Auto Type Casting Manipulation (Node 157) → Line Plot (Node 159) → Image Writer (Port) (Export map)

Recovered Cases till 02-16

# Fitness and Prediction of Virus Cases with R AnalyticFlow

Tool source: <https://r.analyticflow.com/en>

The screenshot displays the R AnalyticFlow interface. The main window shows a plot of 'Cases' vs 'Days from From nCoV-2019 Case First Reported'. The plot includes data points (black crosses), a solid blue line, a dashed red line, and a dashed black curve. The regression equations are:

$$y = -106.714x + 43594 \quad R^2 = 0.943$$
$$y = 21.999x^{1.181} \quad R^2 = 0.927$$
$$y = -0.132x^3 + 4.828x^2 - 1.224x - 27.446 \quad R^2 = 0.943$$

The R Console at the bottom left contains the following code:

```
> lines(x2, y2, lty = 2, col = "black", lwd = 2)
> lines(x2, y6, lty = 2, col = "green", lwd = 2)
> points(x1$ID, x1$value.pch=0, cex=2, col="black")
> text(35.1, labels = equation3.pos=1, adj=0, offset=0, cex=1.2)
> text(35.1, labels = equation2.pos=1, adj=0, offset=-2, cex=1.2)
> text(35.1, labels = equation1.pos=1, adj=0, offset=-4, cex=1.2)
> title(paste(Rname1, "Jan", day2, sep=""))
```

The workflow diagram on the right shows the process: 程序包 (Package) → 数据导入 (Data Import) → 选择城市 (Select City) → 数据整理 (Data Processing) → 简单制图 (Simple Plotting) → Function → plot → prediction → Rname.result2 → Rsquare.

The bottom right section contains a grid of plots for different cities:

- 荆门 Jingmeng Jan2 2
- 荆州 Jingzhou Jan2 2
- 潜江 Qianjiang Jan2 5
- 十堰 Shiyan Jan2 3

Each plot shows '病例数量' (Number of Cases) vs 'Days From nCoV-2019 Case First Reported' with a vertical dashed line indicating the reporting date (e.g., Feb 2nd).

# “The Resources for Novel Coronavirus and Global Research” “新冠病毒资源与全球研究”项目

## ☐ Resource support

- Data
- Tools
- Cloud computation

## ☐ Training

- Data
- Technology
- Methodology
- Applications

## ☐ Distribution

- Research data
- Papers and reports
- Workflows

## ☐ Workshop and symposium



# How Can You Be Involved?

- Make contributions to data collections
- Join research teams
- Share your research results
- Offer training webinars
- Participate in workshops
- Other suggestions?**

For more information, contact: [office@chinadatacenter.net](mailto:office@chinadatacenter.net)

# Acknowledgement

## **Wuhan University**

Xinyan Zhu, Yuanzhen Shao, Lingbo Liu, Yandong Wang, Jing Du, Yanlin Li, Zhiwei Fu, Yicheng Tang, Hongqiang Liu, Huan Zhou, An Li, Lian Duan, Jialei Wang

## **Harvard University**

Wendy Guan, Tao Hu, Wei Luo, Zhibin Li, Danny Brooke

## **RMDS Lab**

Alex Liu, Alicia Wei, Yula Guo, Yuxi Ma, Jidong Li, Mo Chen

## **Vesystem Inc.**

Chong Gao, Ran Wei

# Websites and Contacts

## Resources for COVID-19 Study

<http://chinadatalab.net>

### China Data Online

<http://china-data-online.com>

### China Data Lab on Cloud

<http://chinadatalab.org>

## Contacts

[SpatialDataLab@list.fas.harvard.edu](mailto:SpatialDataLab@list.fas.harvard.edu)

[office@chinadatacenter.net](mailto:office@chinadatacenter.net)

