

The background of the slide is a deep space image featuring a prominent spiral galaxy in the center, surrounded by numerous stars and other distant galaxies. Overlaid on this are several large, semi-transparent, light blue circular shapes that create a layered, abstract effect.

Utilization of Public and Third-Party Data in Understanding Canberra's Population

FIT5137 Advance database
technology

Ziqi Pei
2023-09-13

CONTENTS

01

Slide 1:
Introduction
(30s)

02

Slide 2:
Research
Objectives
(1min)

03

Slide 3:
Leveraging
Public and
Third-Party
Data (1min)

04

Slide 4:
Practical
Examples
(1min)

05

Slide 5:
Benefits of
Data Utilization
(1min)

06

Slide 6: Case
in Point and
Infrastructure
(1min)

01



Slide 1: Introduction (30s)



Slide 1: Introduction (30s)

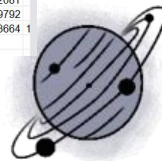
- Title: The Power of Data
- Title: Unveiling Canberra's Population Growth



A. Introduction to the Power of Data

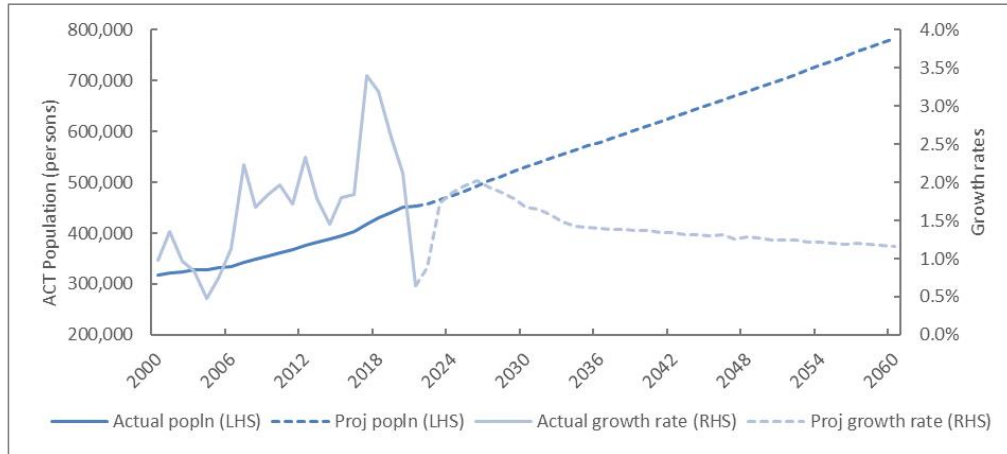
- Data plays a vital role in decision-making, innovation, and research. Public and third-party data are valuable resources for addressing societal, economic, and scientific challenges. This presentation explores the benefits of using such data for research and analyzing demographic shifts in Canberra, the Australian Capital Territory.

1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972
1430972	1443478	1457500	1471066	1492211	1507895	1549199	1602664	1647299	1681469	1703078	1720860	1755325	1789803	1826175	1858002	1892354	1929082	1973700	2004264	2034506	2062692	2096037	2126652	2154753	2186829	2227399	2268665	2373784	2407712
978646	983343	991049	999191	1013867	1030563	1056958	1097556	1136995	1177296	1202936	1231099	1266616	1307253	1338297	1367970	1400192	1437353	1474536	1498161	1526344	1558386	1587580	1614240	1640118	1664208	1693941	1723210	1799486	1826978
540266	545949	552935	559444	567471	580541	595284	615167	631966	647174	662945	676252	692585	708753	723752	735235	749085	763266	779265	789507	802426	818225	834964	849390	861483	875227	891469	905611	932976	956959
304457	307202	310813	314516	320031	329923	340632	357669	369590	381453	392694	403903	415338	430712	442909	453614	464787	477609	491406	496240	509524	522458	537051	550196	556872	562254	570329	579011	597572	605054
245222	247637	250509	252690	258076	264399	273421	286540	298714	309749	320352	330356	338075	346762	352424	357425	363019	367685	380740	390168	401731	412103	421017	432569	449174	468231	489327	508612	539332	553145
122248	123709	125079	126665	129244	132415	136068	140339	145279	151100	155161	157129	159861	162196	165940	169123	172097	174379	177628	179696	182438	184074	185789	187391	189195	191288	193888	195280	199915	200626
7226	7224	7248	7230	7378	8122	8728	9547	10073	9973	10101	10288	11431	12021	12675	13163	14291	15221	25052	25845	27099	28595	29818	31159	34312	37671	40877	44307	47750	50870
7148	7595	8044	8531	9092	11371	12268	13398	13826	14375	15374	16229	17089	18542	20102	22046	24410	27260	30858	34716	38421	41782	46095	49991	53414	57427	62511	67281	71721	81509
3634385	3666337	3730177	3739523	3797370	3865229	3972558	4122880	4253742	4372589	4462641	4546118	4656320	4776042	4882274	4976758	5080235	5192285	5333185	5420867	5522290	5628315	5738351	5841588	5939321	6043235	6170191	6291977	6567936	6685153
1426731	1443096	1468078	1473634	1492627	1507988	1543422	1590708	1630733	1657986	1680714	1702669	1735424	1764454	1796794	1833952	1867480	1903371	1944801	1982684	2015492	2045223	2079401	2112149	2140486	2172396	2213249	2253665	2351719	2387394
997284	1006668	1015980	1025943	1040834	1061757	1086028	1110522	1139580	1167195	1192318	1221242	1250613	1286215	1317960	1350512	1385713	1420036	1455830	1484897	1514497	1547133	1576789	1605977	1634222	1659972	1691102	1721726	1801866	1832276
508151	515947	524189	530113	539844	550659	563840	581018	595735	612304	628465	642007	657432	672838	689333	703984	719152	732661	748249	761475	775441	792473	808570	824934	838499	853769	871618	887132	918509	941519
308592	312241	316493	320313	326042	331230	338673	351878	362840	373900	383087	393191	404229	417945	430257	443189	456111	467711	480081	489256	501417	515562	530520	544788	559557	569004	578976	597602	609574	
230059	233693	237183	238650	244404	250675	258770	270556	281829	290109	300195	309413	319040	327767	335181	342140	349051	354395	368010	375794	386613	396340	404508	415531	430005	446811	465609	482742	514502	528872
120313	121907	123554	125313	127634	128791	130694	135563	140914	145199	148919	151623	154231	156274	160190	163943	167279	169531	172712	175702	178288	180237	182116	184045	186049	188361	191005	192440	198158	199362
3160	3212	3270	3332	3490	4132	4663	5145	5536	5491	5752	6181	6779	7535	8386	9364	9798	10352	19429	20159	21362	22868	24040	25345	27524	29866	32085	34504	37985	41211
6239	6610	6968	7336	7813	8420	9118	10426	11066	11985	13271	14086	15661	16593	17763	19121	21660	24678	27970	31457	34979	38551	42370	46041	50064	54668	59151	64187	74048	78283
3600519	3643374	3688515	3725634	3781988	3843532	3935508	4055816	4168033	4263969	4352721	4440412	4543409	4649521	4757864	4865755	4976244	5082735	5215082	5321424	5428089	5538387	5649314	5757910	5859757	5965400	6092823	6215372	6499329	6618511
2857703	2886574	2918378	2944700	2984838	3015763	3092621	3193372	3278032	3339455	3383792	3423529	3490749	3554257	3624969	3691954	3759834	3832453	3918501	3996948	4049998	4107915	4175438	4237901	4295239	4359325	4441188	4522330	4725503	4795106
1974130	1990011	2007029	2025134	2054701	2092320	2142986	2208078	2276575	2344491	2395254	2452341	2517229	2593468	2656257	2718482	2785905	2857389	2930306	2983058	3040841	3105519	3164369	3220217	3274340	3324180	3380543	3443936	3601352	3661254
1048417	1061896	1077124	1089557	1106415	1131200	1159124	1196185	1227701	1259478	1291410	1318259	1350017	1381591	1413085	1439199	1468237	1495927	1527514	1559082	1577967	1610698	1644534	1674324	1699962	1728996	1763087	1792743	1851485	1898478
613039	619443	627306	634829	646073	661153	679305	709547	732430	755053	775781	797094	819567	848557	873166	898803	920898	945320	971487	987496	1010741	1038020	1067571	1094984	1109780	1121811	1139333	1157987	1200114	1214628
475281	481530	487692	492510	502490	515974	532191	557096	580343	599558	620547	637711	657115	674429	687605	699565	712070	722080	740750	765992	786344	808443	825525	848100	879179	915042	954846	991354	1053834	1082017
242581	245616	248633	251996	257078	261026	267062	275902	286193	296250	304080	308752	314092	318470	325130	330866	339378	343910	355340	355668	360727	364311	367905	371436	375244	379649	384893	387720	396073	400308
10396	10436	10518	10562	10668	12254	13391	14692	15609	15464	15853	16469	18210	19556	21061	22097	24089	25573	44481	46004	48461	51463	53585	56504	61836	67537	72962	78811	85735	92081
13387	14205	15012	15867	16905	19791	21388	23824	24932	26360	28645	30315	32750	35135	37865	41167	46070	52368	58828	66173	73400	80333	88465	96032	103478	112095	121662	131468	151169	159792
7234904	7309711	7391692	7465157	7579358	7708761	7908066	8178696	8421775	8636458	8815362	8986530	9199729	9425563	9640138	9842333	10056479	10275020	10548267	10742291	10950379	11166702	11387665	11599498	11799078	12008635	12263014	12507349	13067265	13303664



B.Unveiling Canberra's Population Growth

- The latest forecasts indicate Canberra's population will grow from 453,558 (June 2021) to 784,043 (June 2060) at an average annual rate of around 1.4%, mainly due to net migration. This has significant implications for urban planning and resource allocation.



02

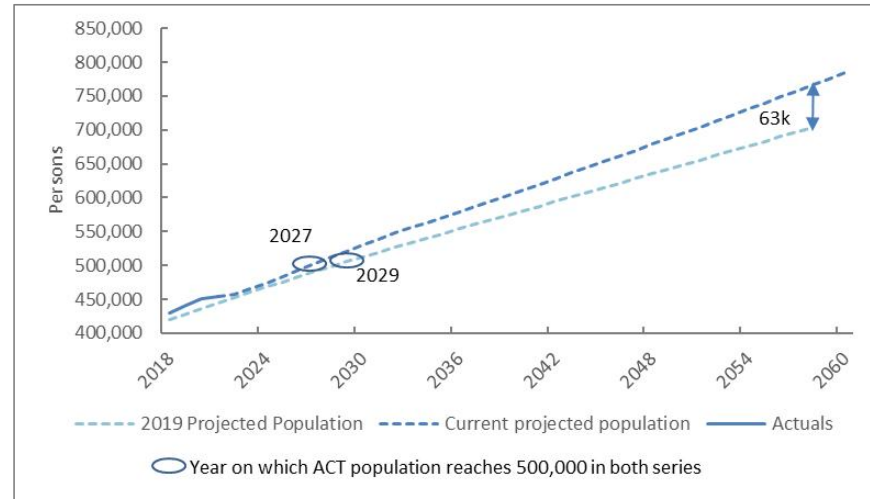


Slide 2: Research Objectives (1min)



A. Broad accessibility:

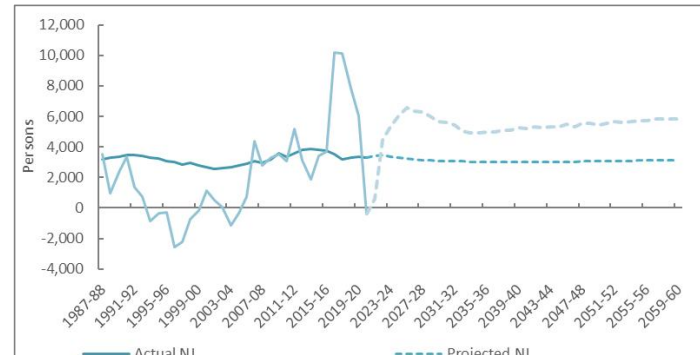
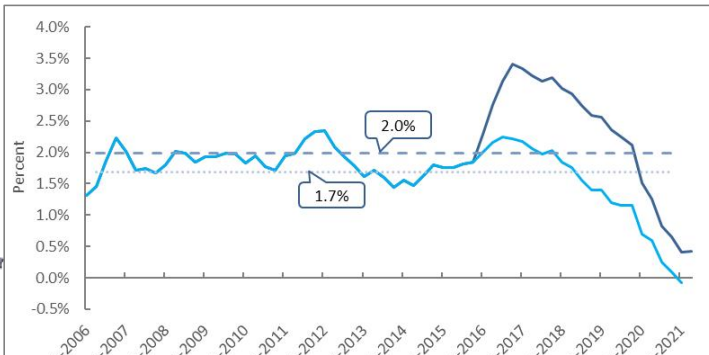
Public data, easily accessible from government and agencies, encourages broad research participation. Figure 2 displays Canberra's population projections, reaching 500,000 despite COVID-19, sourced from the Australian Bureau of Statistics' 2021 Regional Population by Age and Sex data.



B. Historical data and long-term trends:

C. Data consistency and credibility:

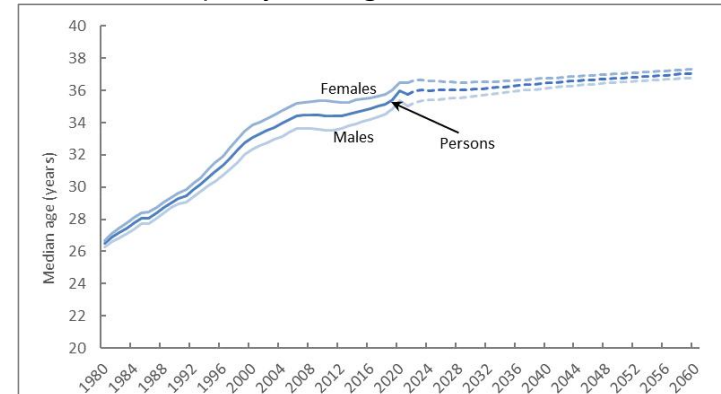
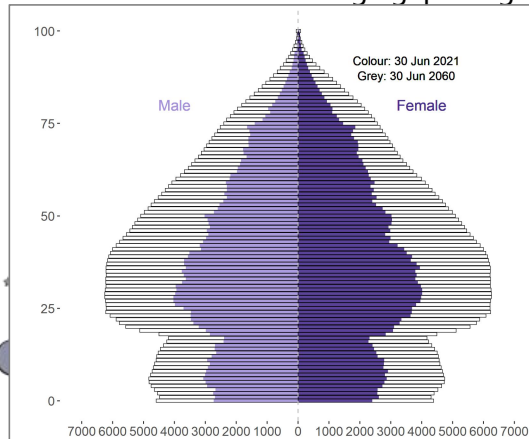
- Public data archives provide crucial historical data for long-term trend analysis. For example, comparing ACT's pre and post-2021 Census population growth rates reveals a significant increase due to the Bureau of Statistics' census adjustments.
- Public data, overseen by government or official bodies, maintain high consistency and trust, alleviating concerns about data quality for researchers.
- ACT's population growth factors include natural growth, net international migration, and net interstate migration. Pre-pandemic, net international migration contributed significantly, but border closures disrupted it. Expected to resume in 2025-26, stabilizing by 2031-32. Limited 2021 Census data; Figure 4 shows historical and future data on natural increase (NI), net migration, including international and interstate migration, and census changes.



D.Cross-field research:

- Public data spans health, education, employment, and the environment, facilitating cross-field research and relationship exploration.
- Age structure: Figure 5 displays ACT's age distribution by gender and age group. While all age groups contribute to population growth from June 2021 to June 2060, there's a noticeable aging trend. The 65+ age group (combined genders) will rise from 13.3% to 14.6%, while the 0-14 age group will decrease from 18.5% to 17.6%. The 15-64 age group, the primary workforce, will see a slight decline.
- Median age: Figure 7 reveals the median age in June 2021 at 35 years, projected to be 37 in June 2060. The gender median age gap will gradually decrease to 37.3 for women and 36.8 for men

are vital for urban planning, resource allocation, and policy making.



03

Slide 3: Leveraging Public and Third-Party Data (1min)



A. Benefits of Public and Third-Party Data in Research

Researchers use two key data types in scientific and social research: public and third-party data. They offer unique advantages for addressing complex issues and informed decision-making.s

SA3	Populati on in June 2021	Populatio n share in June 2021 (%)	Populati on in June 2060	Populatio n share in June 2060 (%)	Change in populati on	Average rate of growth
Belconnen	105,872	23%	175,826	22%	69,954	1.3%
Canberra East	1,932	0%	2,232	0%	300	0.4%
Gungahlin	87,843	19%	148,799	19%	60,956	1.4%
Molonglo	11,382	3%	86,148	11%	74,766	5.3%
North Canberra	61,372	14%	140,999	18%	79,627	2.2%
South Canberra	31,591	7%	58,342	7%	26,751	1.6%
Tuggeranong	89,245	20%	88,914	11%	-331	0.0%
Urriarra – Namadgi	621	0%	469	0%	-152	-0.7%
Weston Creek	24,522	5%	25,671	3%	1,149	0.1%
Woden Valley	39,178	9%	56,643	7%	17,465	0.9%
ACT	452,558	100%	784,042	100%	330,485	1.4%



B. Advantages of Third-Party Data

- **Data Diversity:** Third-party data is diverse, from sources like social media, sensors, satellites, and online transactions, enabling effective problem-solving.
- **Real-Time and High-Resolution Data:** Third-party data offers real-time and high-res info for event monitoring, e.g., social media for sentiment analysis and event tracking.
- **Innovation and Discovery:** Third-party data fuels research innovation with ML, big data, AI for insights & discoveries.
- **Filling Data Gaps:** Third-party data complements public data for fresh perspectives.



C.Utilizing Public and Third-Party Data: Key Advantages

- **Cost-Efficiency:** Public data, often managed by government agencies, reduce data collection costs. Third-party providers streamline data, saving time and expenses.
- **Wide Coverage:** Public data span diverse subjects and regions, ideal for various research. They include census, employment, economy, health, and environmental data.
- **Long-Term Trends:** Public data offer extensive time series, aiding trend analysis, policy-making, and future predictions.
- **Reliability:** Regulated by official agencies, public data maintain high reliability. Quality control and audits ensure data accuracy.
- **Holistic Insights:** Integrating public and third-party data enables comprehensive solutions and deeper insights into complex issues.
- **Real-Time Access:** Some sources provide real-time or near-real-time data, facilitating timely decisions..



04

Slide 4: Practical Examples (1min)



Slide 4: Practical Examples (1min)

- Examples of public and third-party data utilization:
- 1. Outbreak Analysis: During COVID-19, health departments' data aided researchers in analyzing the epidemic's spread, affected regions, and effective interventions.
- 2. Market Research: Analysts use third-party data for market trend analysis, consumer behavior study, and competition assessment. For instance, retail sales data informs product category performance.
- 3. Weather Forecasting: Meteorologists rely on data from weather satellites, ground measurements, and models for accurate predictions. Government and international agencies provide these data.
- 4. Financial Analysis: Financial experts evaluate investment opportunities and market risks using stock market data, financial statements, and macroeconomic data.
- 5. Environmental Monitoring: Environmental scientists use both public and third-party data to monitor air quality, water quality, soil pollution, supporting environmental protection and sustainable development.

• These examples underscore the pivotal role of public and third-party data in diverse fields, empowering decision-



05

Slide 5: Benefits of Data Utilization (1min)



Slide 5: Discuss why utilizing public or third-party data is beneficial for research and analysis and how to reduce data collection costs.(1min)

1. Population Trends: Government census data reveal Canberra's historical population trends, aiding in analyzing growth, decline, and cyclical changes over time.
2. Composition and Characteristics: Public data encompass age, gender, ethnicity, and occupation details, enabling studies on diversity. Third-party sources like social media data provide deeper insights into interests and behaviors.
3. Population Distribution: Public and GIS data help assess population distribution across the city, informing urban planning and resource allocation decisions.
4. Forecasting and Planning: Combining historical public data with third-party insights supports population forecasting and planning for housing, infrastructure, and social services.



Slide 6: Case in Point and Infrastructure (1min)

A. Case in Point: Leveraging Social Media Data for Canberra's Population Study

Analyzing social media data reveals insights into Canberra's population. Residents' online interactions provide information on their interests, activities, and opinions. This data helps identify social trends, event engagement, and sentiment analysis. Monitoring social media discussions informs policy decisions, enhancing urban management and understanding resident reactions to policies and events.

In summary, public and third-party data are vital for grasping Canberra's population dynamics, aiding researchers in understanding the city's evolution. Harnessing these resources effectively yields comprehensive insights for urban planning and policymaking to meet changing societal needs.

B. Building Robust Infrastructure for Data Integration in Population Research

The integration of public and third-party data into population research necessitates a strategic approach:

Collect Data: Gather necessary public and third-party data, such as census, geographic, and economic data, from trustworthy sources for reliability.

Clean and Prepare Data: Thoroughly clean and preprocess data to improve quality. Address missing data, outliers, and format issues. Ensure data consistency across sources.



06

Slide 6: Analyzing Population Data (1min)



A. Case in Point: Leveraging Social Media Data for Canberra's Population Study

Analyzing social media data reveals insights into Canberra's population. Residents' online interactions provide information on their interests, activities, and opinions. This data helps identify social trends, event engagement, and sentiment analysis. Monitoring social media discussions informs policy decisions, enhancing urban management and understanding resident reactions to policies and events.

In summary, public and third-party data are vital for grasping Canberra's population dynamics, aiding researchers in understanding the city's evolution. Harnessing these resources effectively yields comprehensive insights for urban planning and policymaking to meet changing societal needs.

B. Building Robust Infrastructure for Data Integration in Population Research

The integration of public and third-party data into population research necessitates a strategic approach:

Collect Data: Gather necessary public and third-party data, such as census, geographic, and economic data, from trustworthy sources for reliability.

Clean and Prepare Data: Thoroughly clean and preprocess data to improve quality. Address missing data, outliers, and format issues. Ensure data consistency across sources.

Integrate Data: Use suitable techniques to merge data from various sources into a single dataset. Standardize formats and align fields for consistency. Consider database systems or integration tools for efficiency.



Title: Mitigating Risks Associated with External Data

- C.Leveraging Computer Vision for Historical Map Stitching
- Data Collection: Gather high-resolution historical map images from local archives, libraries, and satellite image providers.
- Image Enhancement: Preprocess images to reduce noise and ensure consistent quality by adjusting brightness and contrast.
- Alignment and Registration: Use computer vision techniques to calibrate and align images, identifying reference points for transformation.
- Seamless Stitching: Employ image stitching algorithms, like OpenCV, to merge calibrated images into a continuous historical map.
- Geographical Data Extraction: After stitching, extract geographic information from the map, including details about roads, buildings, and rivers.



Mitigating Risks Associated with External Data Reliance

- **Quality Concerns:** External data may lack rigorous quality control, potentially containing errors that affect analysis.
- **Availability Risks:** Reliance on external sources exposes research or business to interruptions due to service discontinuation or technical issues.
- **Privacy and Legal Hurdles:** Using external data can raise privacy and legal issues, requiring compliance with regulations.
- **Security Threats:** External data sources may contain security vulnerabilities, necessitating strong security measures.
- **Single Source Dependency:** Relying heavily on one external source increases vulnerability to errors or problems.





Slide 7: Conclusion (30s)



Effective Data Management:

- Quality Checks: Prioritize data quality assessments to ensure accuracy.
- Diversify Sources: Use data from multiple sources to reduce reliance on one.
- Data Backup: Establish backup and recovery plans for data availability.
- Regulatory Compliance: Follow data privacy laws and regulations.
- Enhanced Security: Implement comprehensive security measures.
- Consistency: Ensure data consistency when integrating internal and external data.
- Diverse Data Strategy: Incorporate multi-source, open-source, and internal data.
- Sustainability: Plan for long-term data strategies, considering costs and availability.



Title: Conclusion and the Power of Data

- Data is pivotal in comprehending Canberra's population dynamics, aiding in research, decision-making, and policy development. Public and third-party data play vital roles in illuminating urban evolution and societal trends. Let's continue to explore data for deeper insights into our city's future.



- 
- Quote:https://hackerspace.govhack.org/projects/how_canberra_has_evolved_showcasing_historical_imagery
Smith, John. "The Power of Data: How Information Drives Decision Making." Harvard Business Review, 2020.
 - Australian Bureau of Statistics. "Regional Population by Age and Sex, Australia, 2021." ABS, 2021.
 - Brown, Mary, et al. "The Role of Third-Party Data in Market Research." Journal of Marketing Research, vol. 45, no. 6, 2008, pp. 741-758.
 - National Oceanic and Atmospheric Administration (NOAA). "The Importance of Data in Weather Forecasting." NOAA, 2021.
 - Johnson, Robert. "Data-Driven Financial Analysis: Tools and Techniques." Wiley, 2019.
 - Environmental Protection Agency (EPA). "Environmental Data and Monitoring: A Guide to Best Practices." EPA, 2020.
 - World Health Organization (WHO). "Data Utilization in Epidemic Analysis: Lessons from the COVID-19 Pandemic." WHO, 2021.
 - United Nations. "Urban Planning and Sustainable Development: Integrating Population Data." UN, 2018.
 - Zhao, Li, et al. "Enhancing Data Security in the Age of Big Data." IEEE Transactions on Dependable and Secure Computing, vol. 15, no. 4, 2018, pp. 567-578.
 - Government Data Use Agreements: Best Practices and Considerations. Data Governance Institute, 2020.

The Future of Data Collection and Analysis: Trends and Challenges. Data Insights, 2021.