

2018-09

# The 2017 Workshop on Korean Data Services: useful data resources for Korean data reference

Yang, Jungwon

<https://hdl.handle.net/2027.42/145706>

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

---

*Downloaded from Deep Blue, University of Michigan's institutional repository*

# The 2017 Workshop on Korean Data Services: Useful data resources for Korean data reference

Jungwon Yang

International Government Information and Public Policy Librarian  
University of Michigan Library

In North America, scholars' demands for a variety of data resources, such as government statistics, public opinion data, and geo-spatial data, have increased. These scholars are also more likely to apply a variety of research methods, such as qualitative research methods, data visualization, text mining, and other digital humanities techniques. Furthermore, a significant number of academic libraries in North America are developing research data management services to preserve, curate, and disseminate the research data newly generated by these scholars. Given these conditions, East Asian studies librarians have been faced with challenges in providing users with effective data services.

In this paper I will introduce the 2017 Workshop on Korean Data Services that was held at the University of Michigan Library. It was designed to provide useful data service information for Korean Studies librarians and other East Asian Studies librarians as well. I will also explain some useful databases for reference services related to Korean Studies that I introduced in the *International Data class* and the *Korean Data class* at the workshop.

## 1. 2017 Workshop on Korean Data Services

The University of Michigan Library in Ann Arbor, Michigan provided a two-day data services workshop to Korean Studies librarians and other East Asian Studies librarians serving in the United States, Canada, and Europe on November 15 and 16, 2017. The workshop consisted of eleven classes taught by seven subject librarians at the U-M library, who have expertise in international government information, East Asian Studies, digital humanities, data visualization, and research data management. Lynette Hoelter, the Director of Instructional Resources and Development at the Inter-University Consortium for Political and Social Research (ICPSR),<sup>1)</sup> and Hyung Bae Lee, the Korean Studies librarian at Princeton University, participated as instructors.<sup>2)</sup> Twenty-eight academic librarians related to Korean Studies from the United States, Canada, and the Netherlands attended the workshop.

---

1) ICPSR, which was established in 1962, is the largest social science data repository in the United States. It maintains a data archive of more than 250,000 files on research in the social and behavioral sciences. It hosts 21 specialized collections of data on education, aging, criminal justice, substance abuse, terrorism, and other fields. This consortium provides leadership and training in data access, curation, and methods of analysis for the social science research community. ICPSR also provides the *Summer Program in Quantitative Methods of Social Research*, a comprehensive curriculum of intensive courses in research design, statistics, data analysis, and social methodology. <https://www.icpsr.umich.edu/icpsrweb/sumprog/>

2) For more detailed information on the 2017 Workshop on Korean Data Services workshop, please visit the following link: <http://guides.lib.umich.edu/koreadata/workshop2017>



[Figure 1] Participants in the 2017 Workshop on Korean Data Services

In the *International Data* class at this workshop, for example, attendees had a chance to learn how to retrieve data related to Korea, China, and Japan from international governmental organizations' (IGOs) databases, such as the Organisation for Economic Co-operation and Development's (OECD) OECD iLibrary, the International Monetary Fund's (IMF) IMF Data, and the World Bank's World Development Indicator (WDI). In the *Korean Data* class, the instructor provided an overview of the statistics and survey data information from the Korean Statistics Bureau, the Bank of Korea, and the Korean Social Science Data Archive (KOSSDA) at Seoul National University. In these classes, attendees also had opportunities to discuss a range of topics related to data reference services, such as the characteristics and weaknesses/strengths of IGOs' databases, Korean Government agencies' databases, Korean research centers' databases, accession issues for overseas researchers, data selection based on the nature of research, and subject librarian's data consultation services.

## 2. Useful Data for Research related to Korean Studies

### 2.1. International Governmental Organizations' Databases

Thanks to the convenience of internet service, scholars in North America can easily access and acquire data related to foreign countries. According to my personal observations, however, access to "reliable" data resources is not always as easy for them since many online data sources are often inaccurate and outdated. It is difficult to find citation information and ownership of a data set. Given this condition, one of the methods for acquiring "reliable" data sources is to use the data sets of IGOs. The databases from the OECD, IMF, and World Bank provide some good examples.

#### 2.1.1 OECD iLibrary<sup>3)</sup>

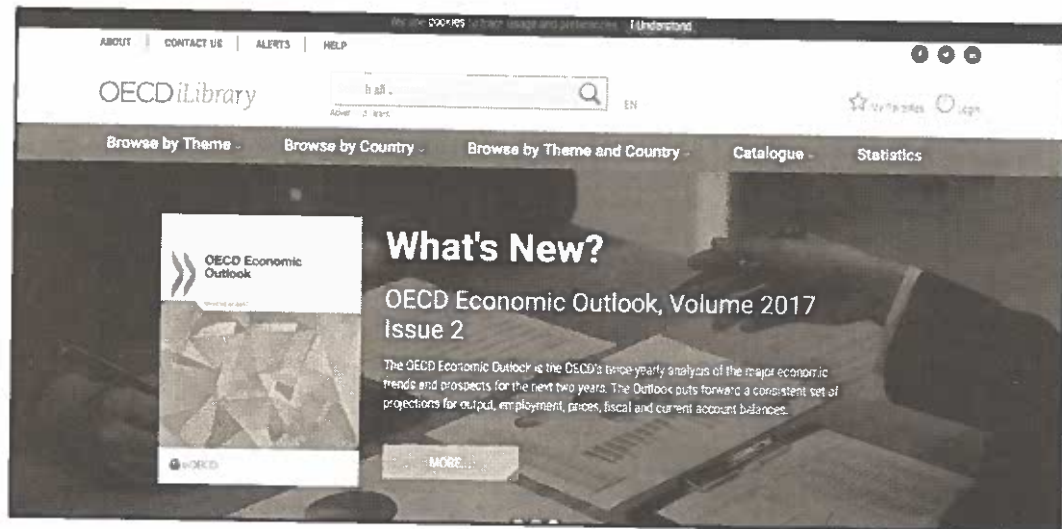
The OECD is an intergovernmental economic organization with 35 member countries. It was founded in 1961 to stimulate economic progress and global trade. It originated in 1948 as the Organization for European Economic Co-operation (OEEC) to allocate United States financial aid and implement economic programs for the reconstruction of Europe following World War II. By the end of the 1950s, with the job of rebuilding Europe effectively complete, OEEC decided to create a new body that would not only assist with European and Atlantic economic issues, but would devise policies to address global economic, political, social, welfare, and environmental issues. South Korea acquired OECD membership in 1996.

The OECD iLibrary is the online library of the OECD and features its books, papers, and statistics. It has served as the gateway to the OECD's analysis and data since 1998. As of January 2018, this database provides 11,710 ebook titles, 175,500 tables and graphs, 5,600 working papers, and 5 billion data points across 42 databases. The seventeen topics which OECD iLibrary currently covers are as follows: economics; education; the environment;

---

<sup>3)</sup> <https://www.oecd-ilibrary.org/>

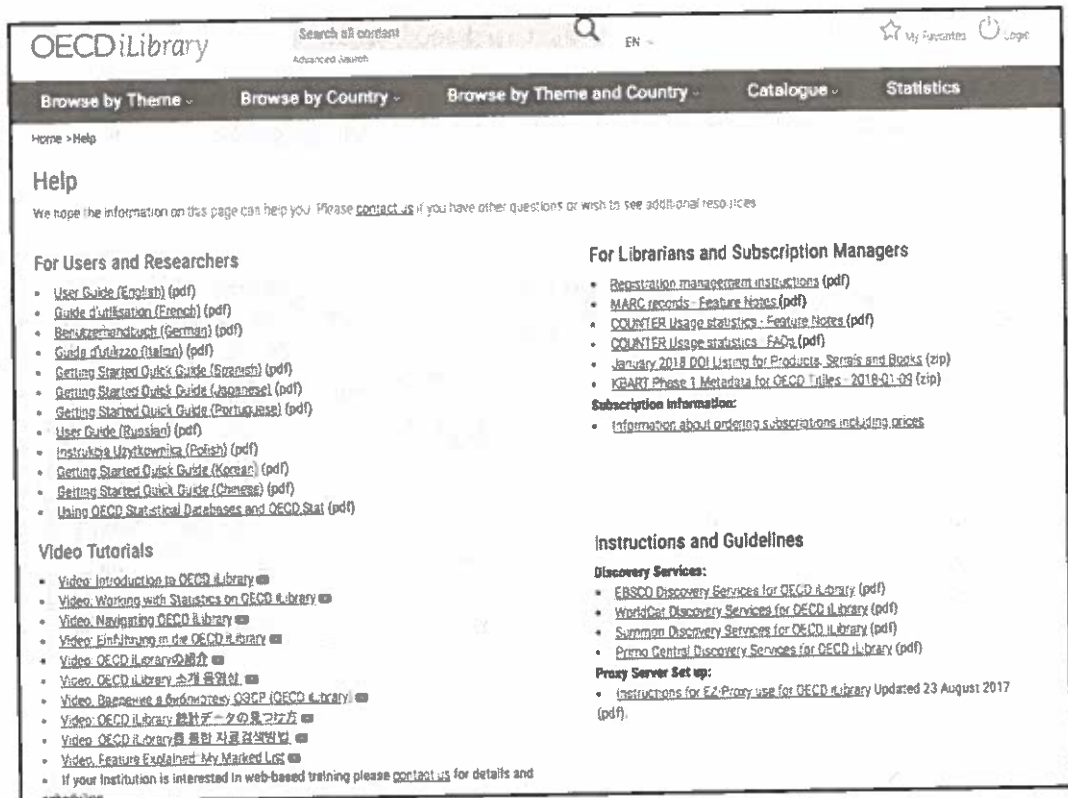
industry and services; transport-ITF; development; finance and investment; social issues/migration/health; urban, rural and regional development; science and technology; energy; governance; taxation; employment; agriculture and food; trade; and nuclear energy.



[Figure 2] OECD iLibrary Homepage

This database was formerly available only to subscribing academic institutions. The OECD recently expanded its free services to the public. For example, ebooks and working papers in the OECD iLibrary that had been available only to subscribers are now accessible in PDF and READ options. Institutions with a subscription can download these ebooks and working papers in PDF format. If the publication contains tables and graphs, this statistical information can be downloaded in Excel and CSV file formats. If a user or an institution has not purchased an OECD iLibrary subscription, they can still access and read the full text of ebooks and working papers through the READ option. They cannot download the full text, graphs, and tables from the database, however. In the case of *OECD iLibrary Statistics*, most datasets (excluding the energy, welfare, and health datasets) are currently open to the public. More detailed information on OECD iLibrary Statistics are available from STATISTICS, which is located in the right upper corner of the OECD iLibrary homepage (please see Figure 2). This also provides user guides and

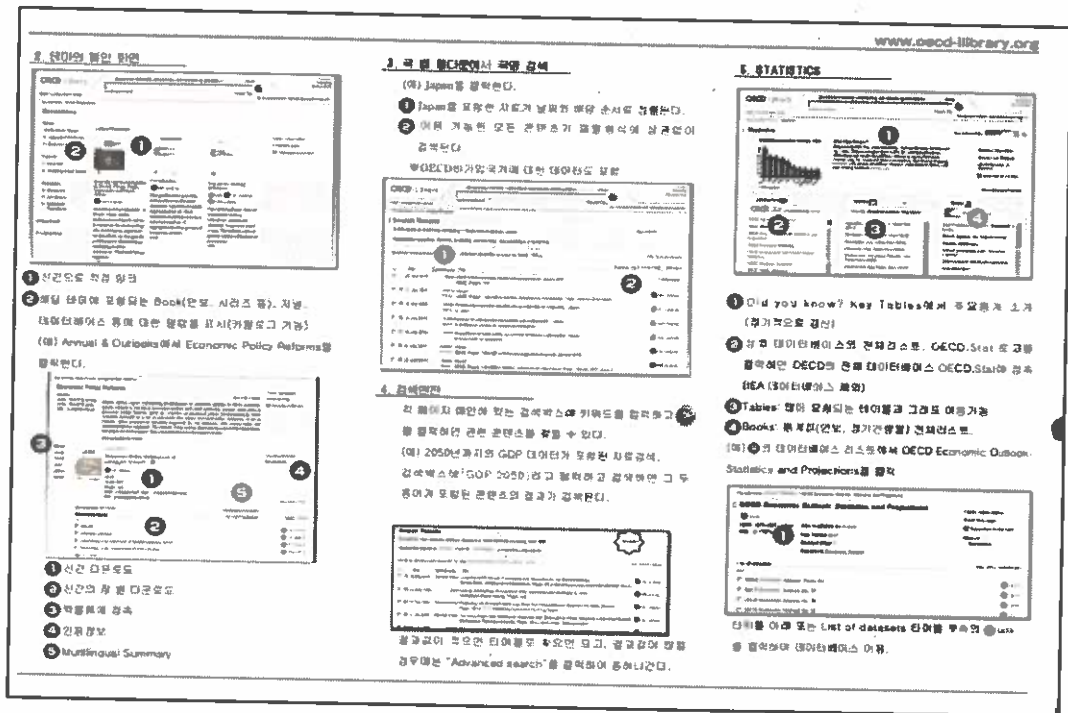
video tutorials in a variety of languages, including Korean.<sup>4)</sup>



[Figure 3] OECD iLibrary Help Page

The Korean-version Pocket Guide, which consists of two pages, and the YouTube video guide are especially useful for Korean Studies librarians who work in a non-English-speaking country.

4) <https://www.oecd-ilibrary.org/userguide/about>



[Figure 4] OECD iLibrary Pocket Guide (Korean Version)

### 2.1.2 IMF Data<sup>5)</sup>

The IMF, also known as the Fund, was conceived in July 1944 at a UN conference in Bretton Woods, New Hampshire in the United States. The 44 countries at the conference sought to build a framework for economic cooperation that could help prevent a repetition of the competitive devaluations that had contributed to the Great Depression of the 1930s. As of 2018, 189 countries are listed as member countries. The IMF Data website is operated by the IMF.

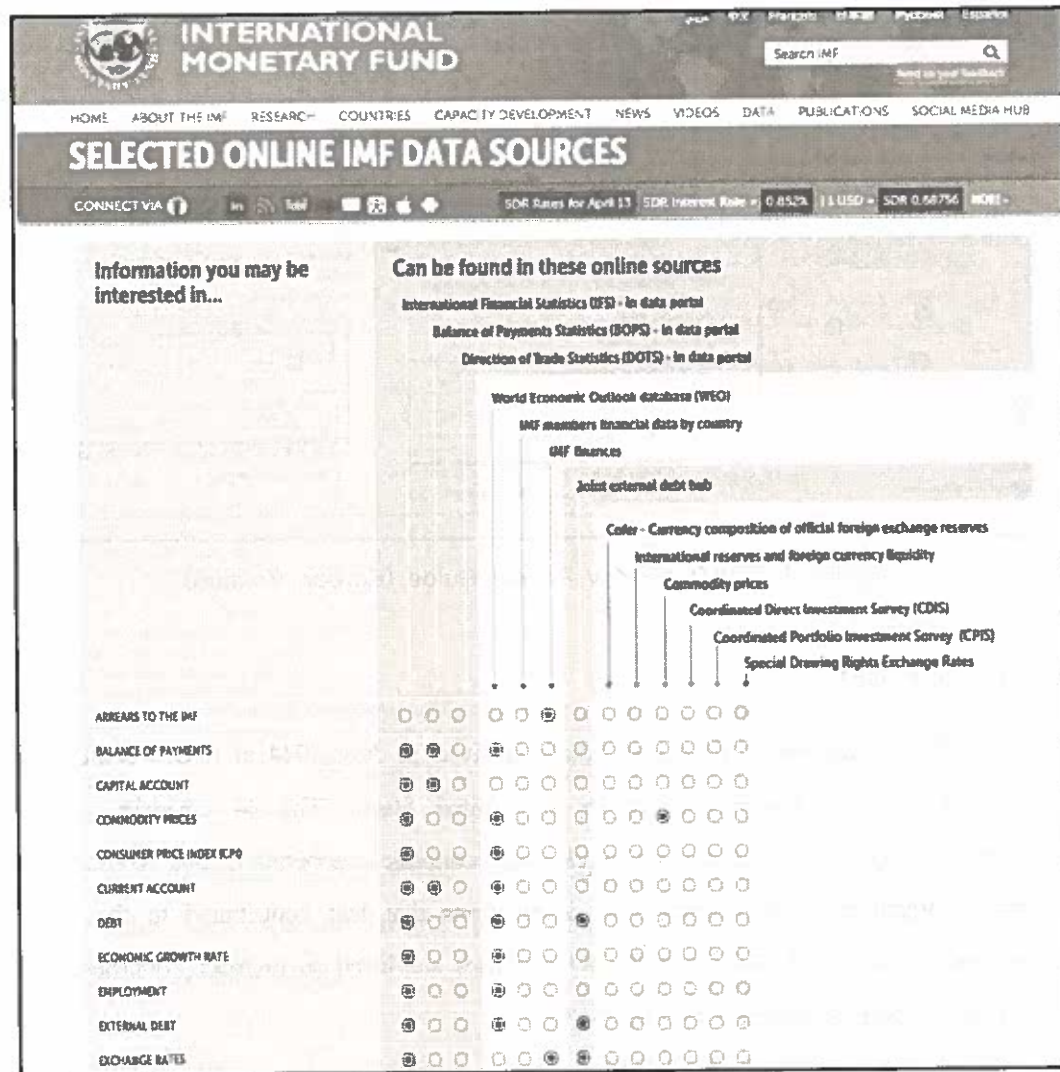
In this database, users can find IMF member countries' macroeconomic and financial data, including exchange rates, IMF lending, balance of payments, and trade. Most of the data housed by IMF Data is open to the public.<sup>6)</sup> As of April 2018, IMF DATA

<sup>5)</sup> <http://www.imf.org/en/Data>

<sup>6)</sup> Except for the Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER), all IMF datasets are open to the public.



provides thirty-nine databases. The information related to these databases can be found on the selected online IMF DATA sources page.<sup>7)</sup>

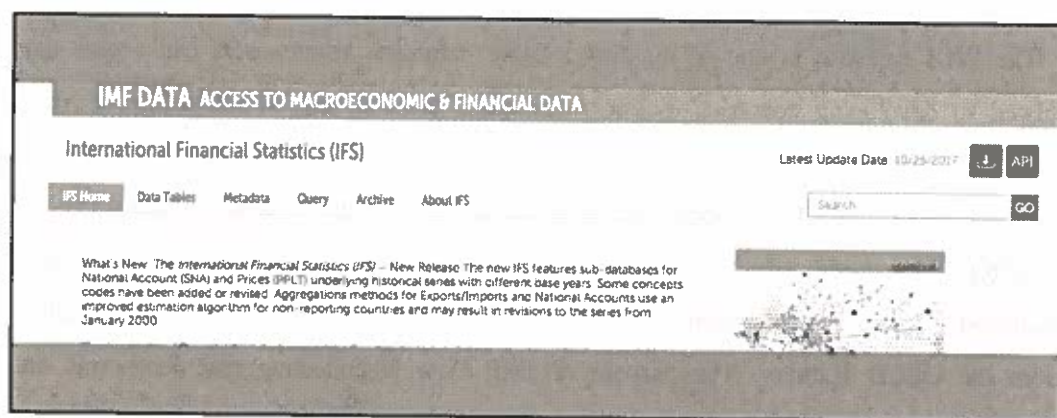


[Figure 5] Selected Online IMF Data Sources

Among these thirty-nine databases, International Financial Statistics (IFS) provides one of the IMF's principal statistical databases providing economic datasets related to Korea

<sup>7)</sup> <http://www.imf.org/external/np/ds/matrix.htm>

since it covers many topics and is focused on data for countries around the world.<sup>8)</sup> Available indicators include a country's exchange rates, fund position, international liquidity, monetary statistics, interest rates, prices, production labor, international transactions, government accounts, national accounts, and population. Some IFS data is available going back to January 1948.



[Figure 6] International Financial Statistics Database

In order to publish data, the Fund regularly collects data from member countries. Based on the data reported by the national authorities, the IMF's Statistics Department (STA) performs certain data transformations to derive higher-level indicators or regional aggregates, as well as cross-country data comparability. STA publishes both data reported by the countries and as calculated by STA in its effort to support the analytical and policy needs of the IMF.

In order to retrieve time-series data from the IFS, users need to use the QUERY option (please see Figure 6). If a user wishes to download a large number of tables from an IMF database or if a user wants to keep the retrieval record, a librarian needs to recommend the creation of an individual account.<sup>9)</sup>

<sup>8)</sup> <http://data.imf.org/?sk=4C514D48-B6BA-49ED-8AB9-52B0C1A0179B>

<sup>9)</sup> The individual account service is provided free of charge.

### 2.1.3 World Development Indicators (World Bank)<sup>10)</sup>

The World Bank, which was organized in August 1946, is an international financial body that provides loans to developing countries in an effort to reduce poverty. The World Development Indicators (WDI) database is one of the World Bank's databases.<sup>11)</sup> The WDI was sold in CD-ROM format until the early 2000s, at which point the World Bank started to provide online service without a subscription fee.<sup>12)</sup>

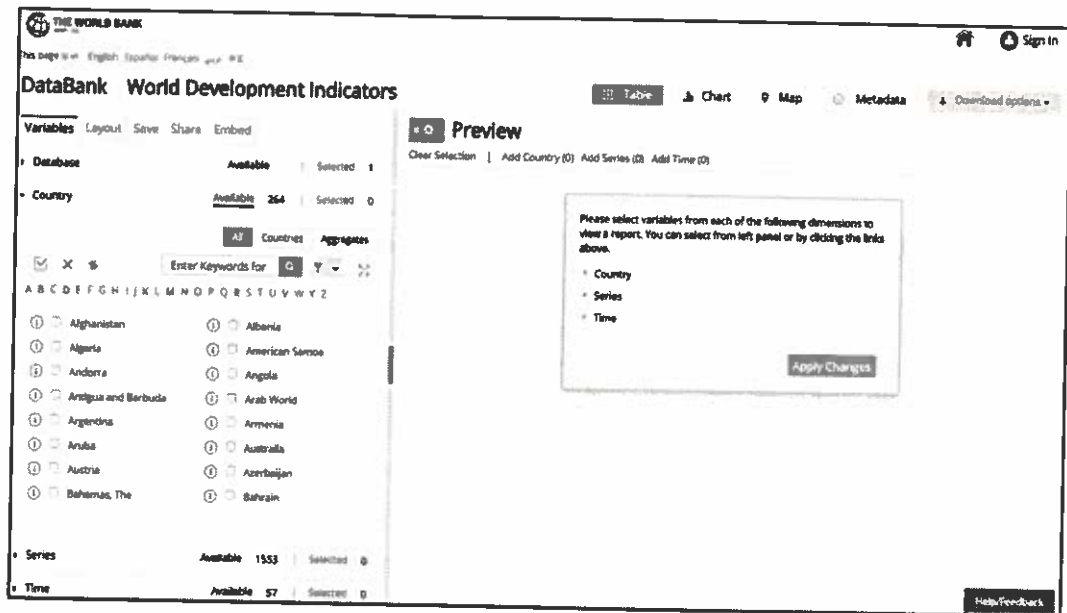
The WDI database is one of the few reliable databases where users can access data related to developing countries' politics, education, environment, social justice, poverty, and other welfare indicators. The OECD iLibrary, for example, focuses on collecting only the data of OECD members, which are relatively wealthy countries. Data related to non-OECD countries, such as Least Developed Countries (LDC), which include most Southeast Asian countries and many African and Latin American nations, is not available from the OECD iLibrary. The purpose of IMF Data is collecting macroeconomic and financial data; therefore, data related to politics, welfare, poverty, and education is not available from this database. The WDI data is available from 1960. All data is yearly so monthly or quarterly data are not available from this database.

---

<sup>10)</sup> <http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators>

<sup>11)</sup> The WDI is a part of the World Bank's Databank, which is an analysis and visualization tool that contains a collection of time series data on a variety of topics. Some examples of useful datasets within the Databank are as follows: World Development Indicators, Statistical Capacity Indicators, Education Statistics, Gender Statistics, and Health Nutrition and Population Statistics. These datasets are all free to use. <http://databank.worldbank.org/data/home.aspx>

<sup>12)</sup> The World Bank eLibrary, which provides ebooks and working papers of the World Bank, is available only to subscribers. <https://elibrary.worldbank.org/>



[Figure 7] The World Bank's WDI Database

Given these characteristics of the databases compiled by IGOs, their strengths are as follows:

The OECD, IMF, and World Bank regularly collect comprehensive information on their members. They also provide member countries with a basic framework for a broader national statistical development strategy and encourage them to use uniform data standards when collecting their data. As a result, data from IGO databases are easy to use for comparative analysis with limited data manipulation. These IGOs also provide in their databases both data collected by member countries and data calculated by their own statistics teams in order to promote the credibility of datasets. Finally, these IGOs tend to expand free data services for users. As a result, users are more likely to access reliable data sets when they use the IGOs' databases.

There are, however, some limitations for these databases. First, international organizations' data is available only as far back as the 1950s. Most of the international organizations were established after the Bretton Woods Conference in 1944 in order to promote global political and economic cooperation. Most European and Asian countries had little time or money to produce government statistics in the 1940s and 1950s after the war.

Scholars who are interested in a period before World War II, therefore, will not find useful data from an IGO database. Second, some monthly and quarterly economic data sets are available from IMF Data and OECD iLibrary databases. However, most of the IGOs' data is yearly data. The IGO databases, therefore, are not a suitable option for research that requires monthly and quarterly data. Finally, it is not a proper resource for research that requires sub-national data sets - provincial, country, township, and city level - because all IGO databases provide only national level data sets. Given these conditions, Korean Studies librarians need to recommend databases from the Korean Statistics Bureau, Bank of Korea, and Korea Social Science Data Archive (KOSIS) for researchers who need pre-1950s data, sub-national data, such as at the city, provincial, county, and township level, and monthly and quarterly data from Korea.

## 2.2 Useful Databases from Korean Government Agencies and Korean Research Centers

### 2.2.1 Korean Statistical Information Services (KOSIS)<sup>13)</sup>

As a gateway for official Korean statistics, KOSIS offers a convenient one-stop service for a full range of major domestic, international, and North Korean data. Currently, it offers official statistics produced by over 300 statistical agencies covering more than 1,000 subjects.

---

<sup>13)</sup> <http://kosis.kr/index/index.do>

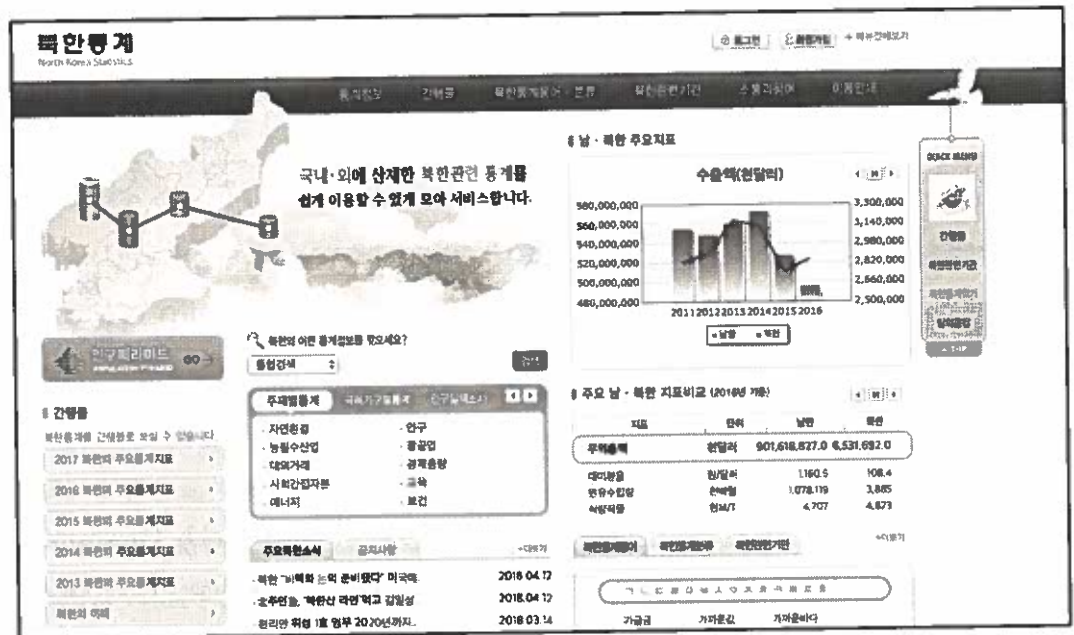


[Figure 8] KOSIS Website (Korean Version)

The KOSIS database encompasses 16 categories: Population/household; Employment/Labor/Wage; Price/Household Income and Expenditure; Health/Society/Welfare; Environment; Agriculture/Forestry and Fishery; Mining and Manufacturing Industry/Energy; Construction/Housing/Land; Transportation/Information and Communication; Wholesale and Retail Trade/Service Industry; Economy/Corporate Business (company); National Accounts/Regional Accounts/National Wealth; Finance/Banking/Insurance; Trade/Foreign Exchange/Balance of Payments; Education/Culture/Science; and Administration. The Korean-language version of KOSIS also provides indicators for local governmental data, past and discontinued statistics databases, pre-colonial period statistics, and North Korea statistics.<sup>14)</sup>

14) <http://kosis.kr/bukhan/index.jsp>

In the North Korea Statistics database, users can access the *Major Statistics Indicators of North Korea*, which is a statistical yearbook.<sup>15)</sup> This publication provides 14 major statistical indicators on North Korea, including population, agriculture/forestry and fishery, international trade, and economy. The publication's data is collected from the United Nations and Korean government agencies. The data is accessible online as well as in PDF file format. All of the North Korea statistical information on this website is open to the public and free of charge.



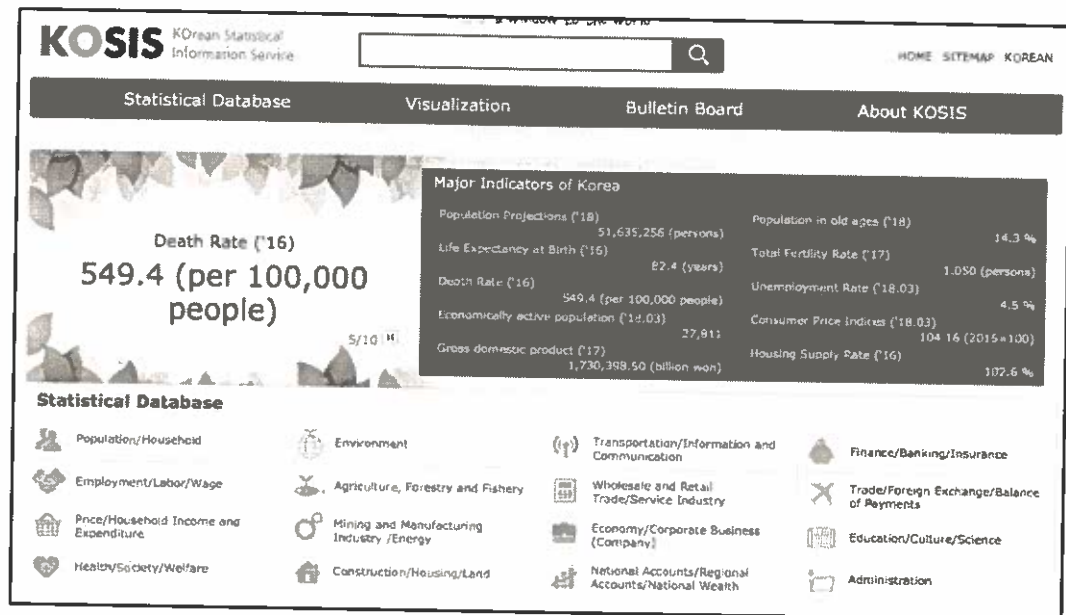
[Figure 9] North Korea Statistics

The KOSIS website, which is governed by the Korean Statistics Bureau, has several strengths compared to other countries' statistics bureau websites. All KOSIS datasets are open to the public and free of charge.<sup>16)</sup> All government agencies' data, local governments'

15) Between 1995 and 2006, the publication of North Korea Statistics was *Comparison between South and North Korea's Social and Economic Status*.

16) In the case of the United States, federal government documents and statistics are open to the public by law. The government information and data sets which are produced by local governments, however, are not. The copyright for local government data, therefore,

statistics, North Korean statistics, pre-1950s data, and discontinued government statistics are available from the KOSIS website. Furthermore, the English version of the KOSIS website is a highly useful resource for researchers who do not understand Korean.



[Figure 10] KOSIS Website (English version)

In order to access the statistics, users need to click the *Statistics Database* tab, which is located in upper left corner of the KOSIS English website (See Figure 10).

There are some limitations of the English version of the website compared to the original Korean website. First of all, pre-1950s data, North Korea statistics, and discontinued data sets are not available from the English version. Second, the metadata of a data set, which is referred to as “statistics information,” is available only from the original website. The statistics information contains information related to data collection period, data source, data format, data creator’s information, and memos from the authors. Overseas Korean Studies librarians, therefore, need to let researchers know that the metadata information

is determined by each local government. In the case of China, government data is not open to the public. The majority of Chinese government statistics can be purchased from vendors.



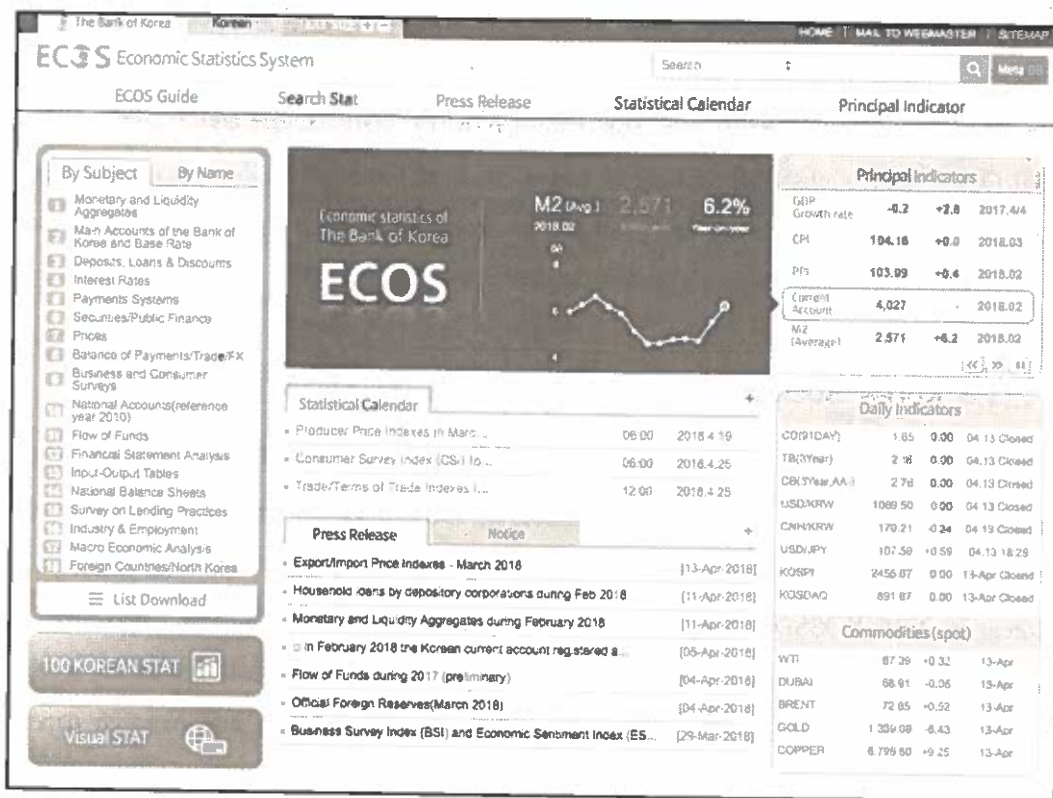
of a dataset is available on the KOSIS original website when they introduce the database to a client.

The screenshot shows a web interface for KOSIS statistics. On the left is a navigation menu titled '주제별통계' (Statistics by Topic) with a sub-section '한국 가구' (Korea Household). The main content area displays a list of statistics, with one entry selected: '성 및 연령별 추계인구(1세별, 5세별) / 전국 수동기'. A pop-up window titled '통계표 정보' (Statistics Information) is overlaid on the right, providing details for the selected statistic.

통계표 정보	
통계표명	성 및 연령별 추계인구(1세별, 5세별) / 전국
수집기간	년 1960-2065
출처	통계청, 광과인구추계 문의치 042-481-7261
분류 및 항목	<input checked="" type="checkbox"/> 가정별 (3개) 중위 추계(기본 추계, 출산율-중위 / 기대수명-중위 / 국제순이동 중위), 고위 추계(최대인구 추계, 출산율-고위 / 기대수명-고위 / 국제순이동-고위), 저위 추계(최소인구 추계, 출산율-저위 / 기대수명-저위 / 국제순이동-저위) <input type="checkbox"/> 성별 (3개) 전체 남자 여자 <input type="checkbox"/> 연령별 (23개) 계 0 - 4세, 0세 1세, 2세, 3세, 4세, 5 - 9세, 5세 6세 <input type="checkbox"/> 항목 (1개) 추계인구
주석	주4) 1960-2015년까지는 확정인구이며, 2015년 이후는 다음

[Figure 11] An Example of Statistics Information

## 2.2.2 The Bank of Korea's Economic Statistical System (ECOS)<sup>17)</sup>



[Figure 12] ECOS (English version)

ECOS is the Bank of Korea's online database. It provides basic economic statistics covering major elements in the Korean economy, such as interest rates, national income, prices, balance of payments, flow of funds, business and consumer surveys, and financial statement analysis. The main users of ECOS are Korean government agencies and policymakers. As of 2014, the ECOS database is open to the public and free of charge. Users can choose either a simple search method or multiple search methods to retrieve data from ECOS. Since the database provides highly granular Korean economic data, it is useful for researchers who want to retrieve daily, monthly, and quarterly Korean

17) <https://ecos.bok.or.kr/EIndex.jsp>

economic and financial data and/or those who need to access more detailed information of economic indicators than the economic data from the KOSIS database. For example, the monthly prices of certain fruits in 1971 is available from KOSIS. The monthly prices of apples, bananas, and oranges in 1971, however, is not available from KOSIS. This data can only be found in ECOS. ECOS provides both Korean and English versions.<sup>18)</sup> Other publications of the Bank of Korea, such as the Monthly Survey Statistics and Economic Statistics Yearbook, are available to download in both Excel and PDF formats on the ECOS website.

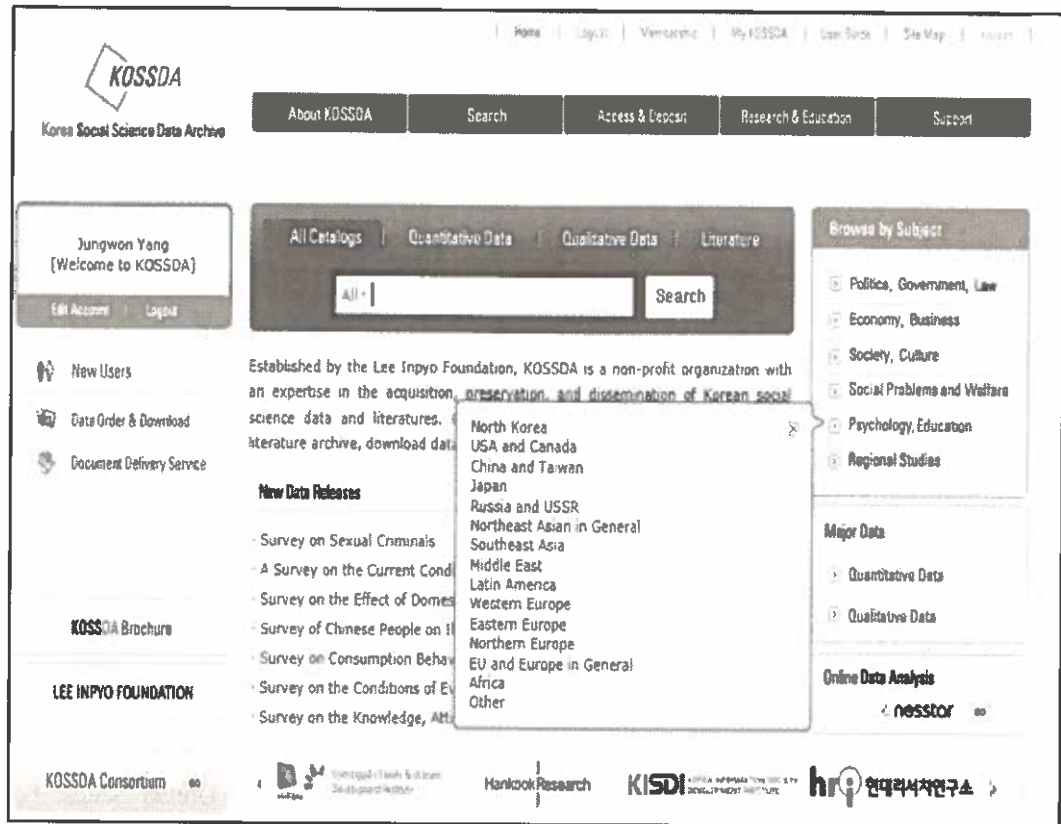
### 2.3.3 Korea Social Science Data Archive (KOSSDA)

KOSSDA is a non-profit organization for Korean social science data services, which was established by the Lee Inpyo Foundation in 1983. It is the first data archive which provides research data acquisition, curation, preservation, and dissemination service in Korea. In 2016 KOSSDA merged with the Asia Center at Seoul National University. As of 2018, 27 Korean government-affiliated research centers and more than 100 social science researchers voluntarily deposit their data, statistics, working papers, and research papers.

KOSSDA holds in its collection a vast range of Korean quantitative and qualitative data and literature across diverse social science disciplines encompassing political, economic, social, and cultural fields. The quantitative data collection includes data from more than 1,900 surveys and aggregated statistical data, such as the Korean General Social Survey (KGSS) and International Social Survey Programme (ISSP) data. In the case of qualitative data collection, more than 200 documents, field notes, interviews, and narrative history records related to regional communities, political and social movements, poverty, foreign workers, and immigration are available through this database. Both English and Korean versions of KOSSDA are available for overseas users.

---

<sup>18)</sup> [https://ecos.bok.or.kr/EIndex\\_en.jsp](https://ecos.bok.or.kr/EIndex_en.jsp)



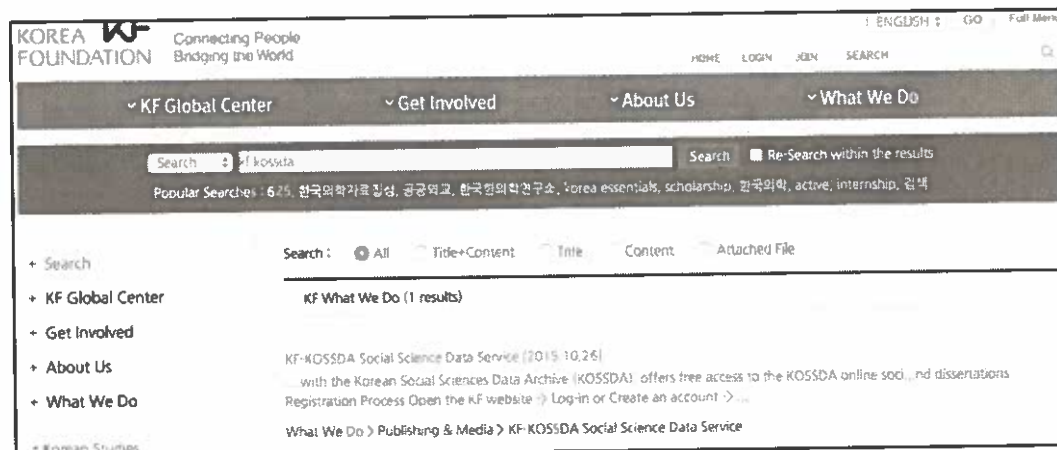
[Figure 13] KOSSDA Website (English version)

The KOSSDA database is a subscription-based online database for domestic users. Overseas researchers, however, can access the database free of charge. In order to use the free service, researchers must first of all create an individual account for the Korean Foundation (KF) website.<sup>19)</sup> After a researcher's individual KF account is approved, the researcher can submit an application for KOSSDA from the KF-KOSSDA Social Science Data Service webpage.<sup>20)</sup> <sup>21)</sup> After an overseas user earns approval from the KF-KOSSDA, users can go to the KF-KOSSDA link to access the KOSSDA database whenever they need.

19) <https://en.kf.or.kr>

20) Applicants need to provide an email address affiliated with an academic institution when they submit the KOSSDA applications.

21) <http://en.kf.or.kr/?menu=3830>



[Figure 14] Kf-KOSSDA Social Science Data Service

### 3. Conclusion

Finding reliable and accessible statistical data for research is not always an easy task for subject librarians and area specialists at academic libraries in the United States. For example, one main purpose of the IGOs' databases is helping the policy-making process at their organizations. They tend to collect, preserve, and disseminate data related to the organization's interests. As a result, it is quite difficult to find an IGO database that covers overall political, social, economic, and scientific topics. IGOs' databases are also not a suitable option for researchers who focus on the pre-World War II period and sub-national regions in Korea. Librarians, therefore, should recommend the use of the KOSIS, ECOS, and KOSSDA databases instead of IGOs' databases for researchers who wish to acquire historic and sub-national level data on Korea. If a researcher is not familiar with the Korean language and requires data sets which they can easily acquire from IGOs' databases, librarians and area experts should not recommend the Korean version of the KOSIS, KOSSDA, or ECOS databases as primary data resources. I, therefore, suggest that Korean studies librarians have a comprehensive data consultation meeting with researchers before they recommend databases to them in order to identify which

database is suitable for their research.

While overseas researchers' interests in Korean data have significantly increased, only a small number of Korean government agencies and Korean research institutions' datasets are translated into English. It is very difficult for overseas researchers to find new datasets from these organizations by themselves. Even if they find useful datasets from Korean government agencies and Korean research institutions, they often give up on using these datasets because of the language barrier, or sometimes misinterpret the data. The role of overseas Korean studies librarians, therefore, is key in helping overseas researchers' efforts. I hope the database information of IGOs, Korean government agencies, and Korean academic institutions that I have introduced in this article is useful for overseas researchers and Korean Studies librarians as well.

task  
For  
cess  
the  
that  
also  
and  
the  
who  
not  
quire  
rean  
fore,  
ting  
hich