Minneapolis College Library Springshare Data Analysis

Asset Views

William Vann

Information Studies Faculty

Summer 2023

An "asset" is a library resource that's linked to in the SpringShare ecosystem. There are 4 categories of assets:

1. ☑ Book
2. ☑ Link
3. ☑ Database
4. Document

This analysis will concern itself with 1-3 only.

In [1]:
```python
# Import dependencies

from datetime import datetime as dt
from scipy import stats as st

import matplotlib.dates as mdates
import matplotlib.pyplot as plt

import numpy as np
import pandas as pd
```

In [2]:
```python
def num_wanted():
    try:
        top_num_wanted = int(input("Enter number wanted: "))

    except:
        print("Incorrect input ... provide an integer.")

    return top_num_wanted
```

In [3]:
```python
assets_df = pd.read_csv("assets_alltime.csv")
assets_df = assets_df[["ID", "Name", "Type", "Total"]]
assets_df = assets_df.loc[(assets_df["Name"] != "[Deleted]"), :]

assets_df = assets_df.sort_values(by="Total", ascending=False)
assets_df = assets_df.reset_index(drop=True)
assets_df
```

Out[3]:

	ID	Name	Type	Total
0	3361065	Academic Search Premier (EBSCO)	Database	25942
1	3361083	CINAHL Ultimate (EBSCO)	Database	6432
2	3361103	Gale Ebooks (Virtual Reference Library)	Database	5387
3	3361091	CQ Researcher	Database	4319
4	3361124	MasterFILE Complete (EBSCO)	Database	3576
...	...	...	...	...
1143	70067705	National Archives: Analyze a Photograph	Link	1
1144	8295358	NIDAMED Clinical Resources	Link	1
1145	7038461	More FRONTLINE	Link	1
1146	50906115	Minnesota.gov Portal	Link	1
1147	69206046	"Anarchism." Encyclopedia of Activis...	Link	1

1148 rows × 4 columns

In [4]:
```python
assets_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1148 entries, 0 to 1147
Data columns (total 4 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   ID      1148 non-null   int64
 1   Name    1148 non-null   object
 2   Type    1148 non-null   object
 3   Total   1148 non-null   int64
dtypes: int64(2), object(2)
memory usage: 36.0+ KB
```

In [5]:
```python
assets_total = assets_df.shape[0]
highest_views = assets_df.iloc[0, -1]
lowest_views = assets_df.iloc[-1, -1]

print(f"{assets_total} assets total, ranging from {lowest_views} to {highest_views} accesses."
```

```
1148 assets total, ranging from 1 to 25942 accesses.
```

In [6]:
```python
num_leaders = num_wanted()

leaders = assets_df.sort_values(by="Total", ascending=False).head(num_leaders)
leaders = leaders.sort_values(by="Total", ascending=True)

leaders = leaders.reset_index(drop=True)

leaders
```
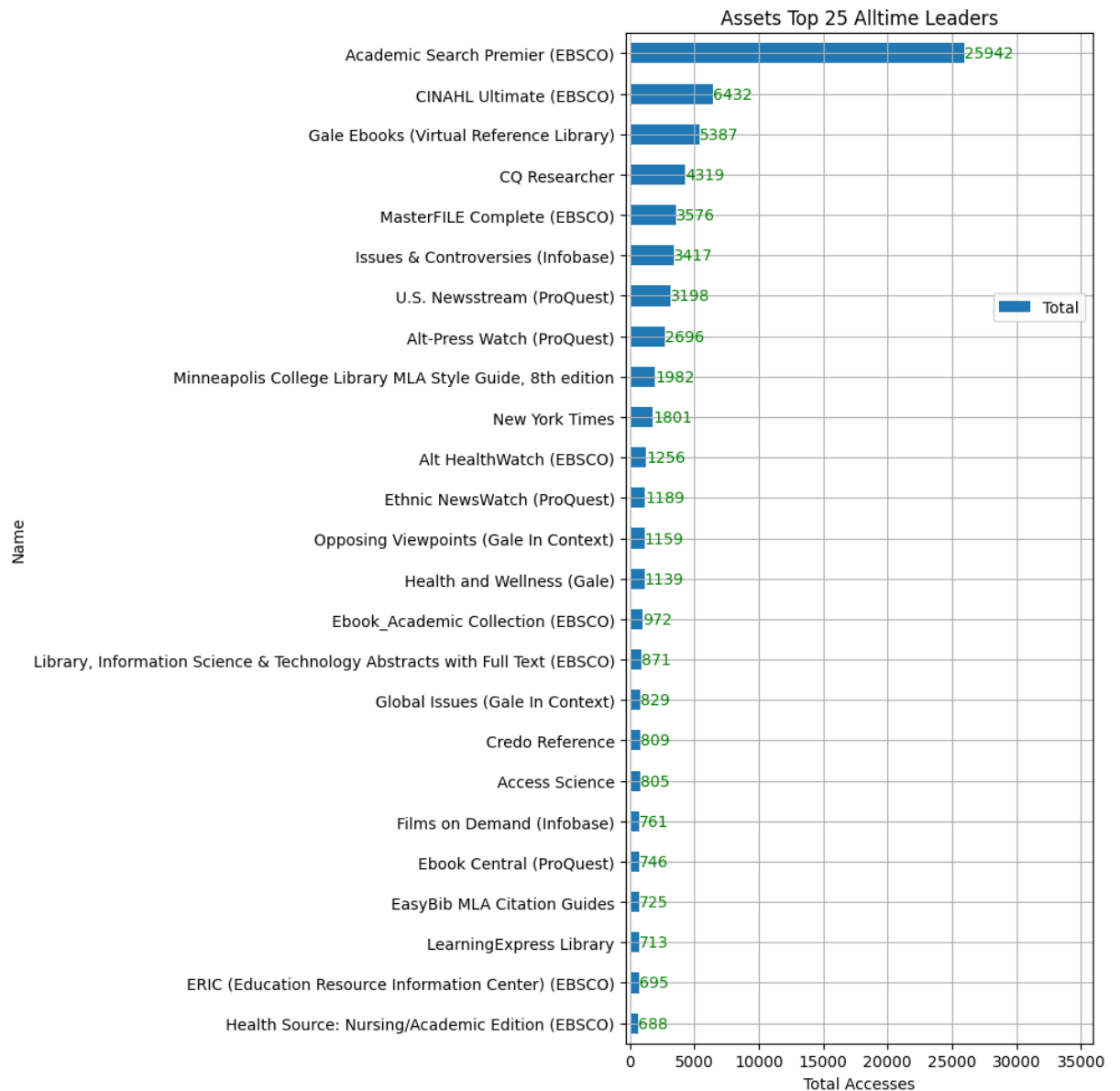
Enter number wanted: 25

Out[6]:

	ID	Name	Type	Total
0	3361112	Health Source: Nursing/Academic Edition (EBSCO)	Database	688
1	3361098	ERIC (Education Resource Information Center) (...	Database	695
2	10410784	LearningExpress Library	Database	713
3	20439402	EasyBib MLA Citation Guides	Link	725
4	3361095	Ebook Central (ProQuest)	Database	746
5	3361102	Films on Demand (Infobase)	Database	761
6	3361066	Access Science	Database	805
7	3361093	Credo Reference	Database	809
8	3361106	Global Issues (Gale In Context)	Database	829
9	3361121	Library, Information Science & Technology Abst...	Database	871
10	3361163	Ebook_Academic Collection (EBSCO)	Database	972
11	3361109	Health and Wellness (Gale)	Database	1139
12	3361135	Opposing Viewpoints (Gale In Context)	Database	1159
13	3361099	Ethnic NewsWatch (ProQuest)	Database	1189
14	3361068	Alt HealthWatch (EBSCO)	Database	1256
15	22812778	New York Times	Database	1801
16	24834464	Minneapolis College Library MLA Style Guide, 8...	Link	1982
17	3361071	Alt-Press Watch (ProQuest)	Database	2696
18	3361145	U.S. Newsstream (ProQuest)	Database	3198
19	3361115	Issues & Controversies (Infobase)	Database	3417
20	3361124	MasterFILE Complete (EBSCO)	Database	3576
21	3361091	CQ Researcher	Database	4319
22	3361103	Gale Ebooks (Virtual Reference Library)	Database	5387
23	3361083	CINAHL Ultimate (EBSCO)	Database	6432
24	3361065	Academic Search Premier (EBSCO)	Database	25942

In [7]:

```python
ax = leaders.plot(
        x="Name",
        y="Total",
        figsize=(10,10),
        kind="barh",
        grid=True,
        xlim=[leaders["Total"].min()-1000, leaders["Total"].max()+10000],
        xlabel="Total Accesses",
        title=f"Assets Top {num_leaders} Alltime Leaders")

ax.bar_label(ax.containers[0], color="green")
ax.legend(loc="best", bbox_to_anchor=(0.5, 0.25, 0.5, 0.5))
plt.tight_layout()

plt.savefig(f"Assets/assets_{num_leaders}.png")

plt.show()
```

Assets Top 25 Alltime Leaders

Name	Total
Academic Search Premier (EBSCO)	25942
CINAHL Ultimate (EBSCO)	6432
Gale Ebooks (Virtual Reference Library)	5387
CQ Researcher	4319
MasterFILE Complete (EBSCO)	3576
Issues & Controversies (Infobase)	3417
U.S. Newsstream (ProQuest)	3198
Alt-Press Watch (ProQuest)	2696
Minneapolis College Library MLA Style Guide, 8th edition	1982
New York Times	1801
Alt HealthWatch (EBSCO)	1256
Ethnic NewsWatch (ProQuest)	1189
Opposing Viewpoints (Gale In Context)	1159
Health and Wellness (Gale)	1139
Ebook_Academic Collection (EBSCO)	972
Library, Information Science & Technology Abstracts with Full Text (EBSCO)	871
Global Issues (Gale In Context)	829
Credo Reference	809
Access Science	805
Films on Demand (Infobase)	761
Ebook Central (ProQuest)	746
EasyBib MLA Citation Guides	725
LearningExpress Library	713
ERIC (Education Resource Information Center) (EBSCO)	695
Health Source: Nursing/Academic Edition (EBSCO)	688

In [8]:
```python
num_leaders = num_wanted()

book_leaders = assets_df.loc[(assets_df["Type"]=="Book"), :].sort_values(by="Total", ascending
book_leaders = book_leaders.sort_values(by="Total", ascending=True).reset_index(drop=True)

book_leaders
```
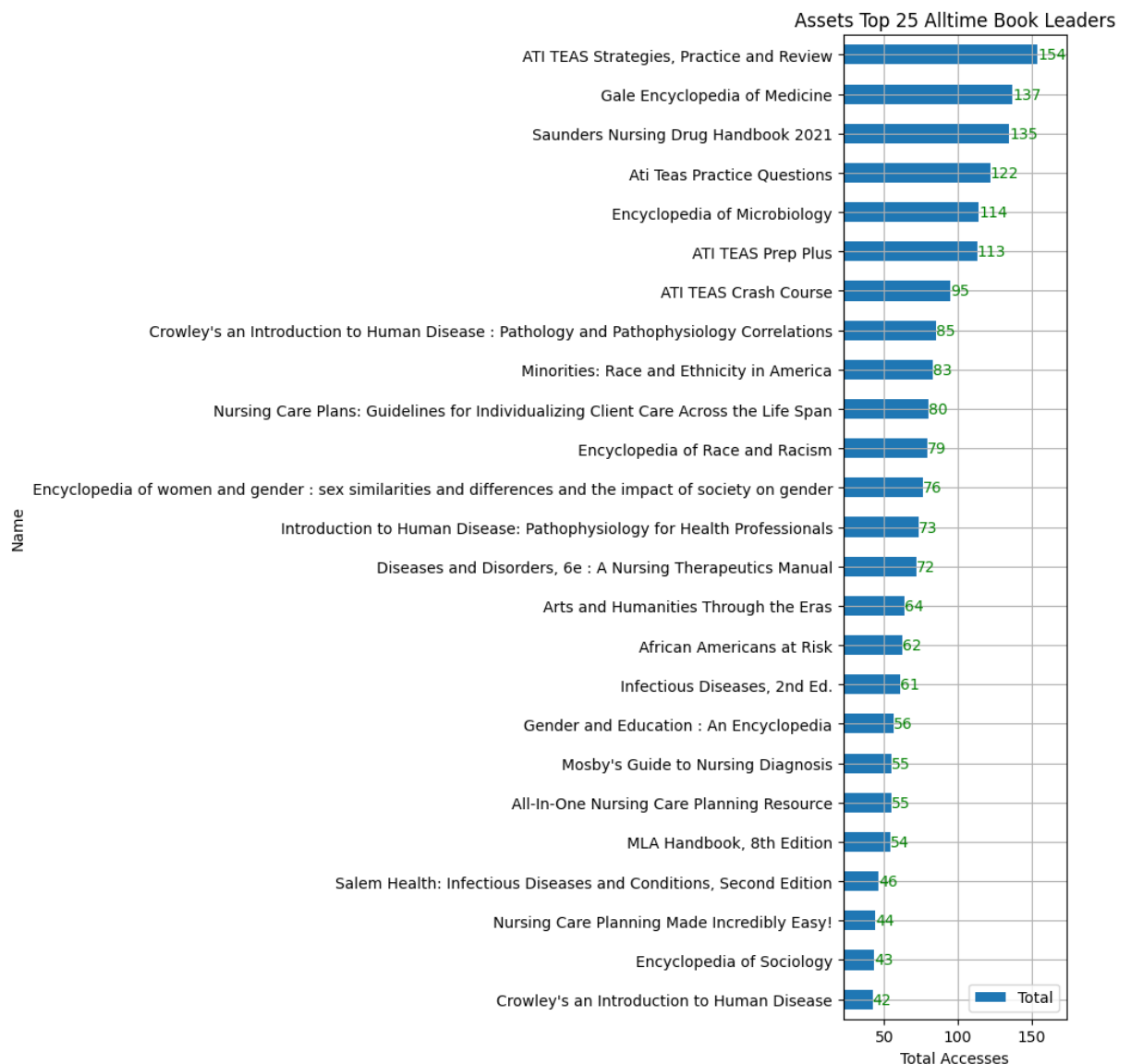
Enter number wanted: 25

Out[8]:

	ID	Name	Type	Total
0	8425558	Crowley's an Introduction to Human Disease	Book	42
1	4955912	Encyclopedia of Sociology	Book	43
2	57917020	Nursing Care Planning Made Incredibly Easy!	Book	44
3	3361684	Salem Health: Infectious Diseases and Conditio...	Book	46
4	20407948	MLA Handbook, 8th Edition	Book	54
5	24974234	All-In-One Nursing Care Planning Resource	Book	55
6	55570344	Mosby's Guide to Nursing Diagnosis	Book	55
7	5090986	Gender and Education : An Encyclopedia	Book	56
8	3361686	Infectious Diseases, 2nd Ed.	Book	61
9	46366881	African Americans at Risk	Book	62
10	3361783	Arts and Humanities Through the Eras	Book	64
11	61725848	Diseases and Disorders, 6e : A Nursing Therape...	Book	72
12	55570659	Introduction to Human Disease: Pathophysiology...	Book	73
13	5091030	Encyclopedia of women and gender : sex similar...	Book	76
14	28982546	Encyclopedia of Race and Racism	Book	79
15	62292898	Nursing Care Plans: Guidelines for Individuali...	Book	80
16	46366880	Minorities: Race and Ethnicity in America	Book	83
17	55570696	Crowley's an Introduction to Human Disease : P...	Book	85
18	12342734	ATI TEAS Crash Course	Book	95
19	12342257	ATI TEAS Prep Plus	Book	113
20	59247517	Encyclopedia of Microbiology	Book	114
21	12342258	Ati Teas Practice Questions	Book	122
22	55570607	Saunders Nursing Drug Handbook 2021	Book	135
23	3361679	Gale Encyclopedia of Medicine	Book	137
24	41212333	ATI TEAS Strategies, Practice and Review	Book	154

In [9]:

```python
ax = book_leaders.plot(
        x="Name",
        y="Total",
        figsize=(10,10),
        kind="barh",
        grid=True,
        xlim=[book_leaders["Total"].min()-20, book_leaders["Total"].max()+20],
        xlabel="Total Accesses",
        title=f"Assets Top {num_leaders} Alltime Book Leaders")

ax.bar_label(ax.containers[0], color="green")
plt.tight_layout()

plt.savefig(f"Assets/book_assets_{num_leaders}.png")

plt.show()
```


Assets Top 25 Alltime Book Leaders

In [10]:

```python
num_leaders = num_wanted()

link_leaders = assets_df.loc[(assets_df["Type"]=="Link"), :].sort_values(by="Total", ascending
link_leaders = link_leaders.sort_values(by="Total", ascending=True).reset_index(drop=True)

link_leaders
```

Enter number wanted: 25

Out[10]:

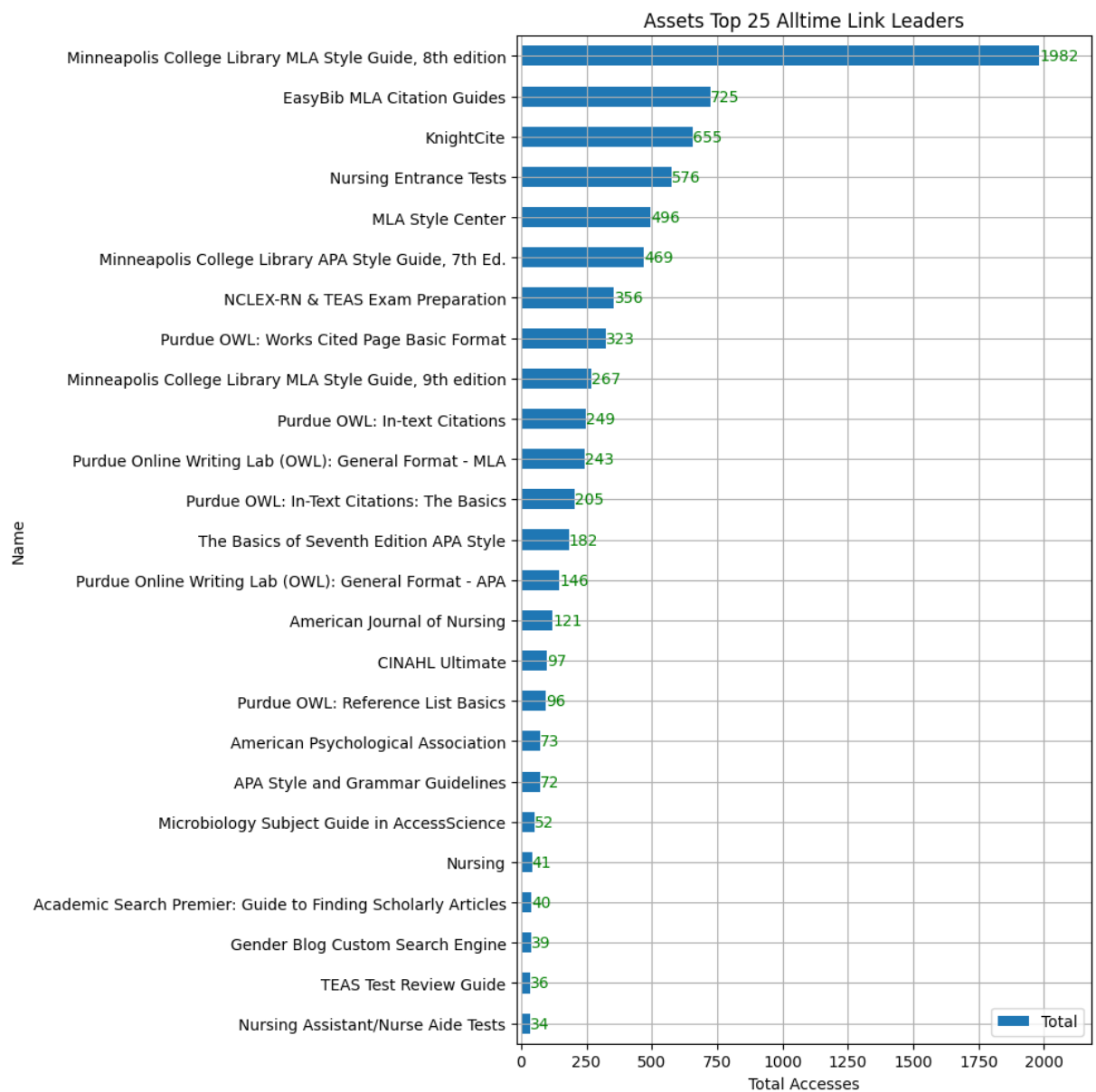|	ID	Name	Type	Total
0	54933110	Nursing Assistant/Nurse Aide Tests	Link	34
1	66379075	TEAS Test Review Guide	Link	36
2	46841420	Gender Blog Custom Search Engine	Link	39
3	3361938	Academic Search Premier: Guide to Finding Scho...	Link	40
4	18970578	Nursing	Link	41
5	50529074	Microbiology Subject Guide in AccessScience	Link	52
6	20407985	APA Style and Grammar Guidelines	Link	72
7	20407983	American Psychological Association	Link	73
8	20407991	Purdue OWL: Reference List Basics	Link	96
9	39844473	CINAHL Ultimate	Link	97
10	6264134	American Journal of Nursing	Link	121
11	49828623	Purdue Online Writing Lab (OWL): General Forma...	Link	146
12	20407984	The Basics of Seventh Edition APA Style	Link	182
13	20407989	Purdue OWL: In-Text Citations: The Basics	Link	205
14	20407956	Purdue Online Writing Lab (OWL): General Forma...	Link	243
15	20407958	Purdue OWL: In-text Citations	Link	249
16	68218206	Minneapolis College Library MLA Style Guide, 9...	Link	267
17	20407962	Purdue OWL: Works Cited Page Basic Format	Link	323
18	10410236	NCLEX-RN & TEAS Exam Preparation	Link	356
19	3362128	Minneapolis College Library APA Style Guide, 7...	Link	469
20	40113859	MLA Style Center	Link	496
21	10409988	Nursing Entrance Tests	Link	576
22	40113918	KnightCite	Link	655
23	20439402	EasyBib MLA Citation Guides	Link	725
24	24834464	Minneapolis College Library MLA Style Guide, 8...	Link	1982

In [11]:

```python
ax = link_leaders.plot(
        x="Name",
        y="Total",
        figsize=(10,10),
        kind="barh",
        grid=True,
        xlim=[link_leaders["Total"].min()-50, link_leaders["Total"].max()+200],
        xlabel="Total Accesses",
        title=f"Assets Top {num_leaders} Alltime Link Leaders")

ax.bar_label(ax.containers[0], color="green")
plt.tight_layout()

plt.savefig(f"Assets/link_assets_{num_leaders}.png")

plt.show()
```

Assets Top 25 Alltime Link Leaders

Name	Total
Minneapolis College Library MLA Style Guide, 8th edition	1982
EasyBib MLA Citation Guides	725
KnightCite	655
Nursing Entrance Tests	576
MLA Style Center	496
Minneapolis College Library APA Style Guide, 7th Ed.	469
NCLEX-RN & TEAS Exam Preparation	356
Purdue OWL: Works Cited Page Basic Format	323
Minneapolis College Library MLA Style Guide, 9th edition	267
Purdue OWL: In-text Citations	249
Purdue Online Writing Lab (OWL): General Format - MLA	243
Purdue OWL: In-Text Citations: The Basics	205
The Basics of Seventh Edition APA Style	182
Purdue Online Writing Lab (OWL): General Format - APA	146
American Journal of Nursing	121
CINAHL Ultimate	97
Purdue OWL: Reference List Basics	96
American Psychological Association	73
APA Style and Grammar Guidelines	72
Microbiology Subject Guide in AccessScience	52
Nursing	41
Academic Search Premier: Guide to Finding Scholarly Articles	40
Gender Blog Custom Search Engine	39
TEAS Test Review Guide	36
Nursing Assistant/Nurse Aide Tests	34

In [12]: ▶
```
num_leaders = num_wanted()

database_leaders = assets_df.loc[(assets_df["Type"]=="Database"), :].sort_values(by="Total", a
database_leaders = database_leaders.sort_values(by="Total", ascending=True).reset_index(drop=T

database_leaders
```
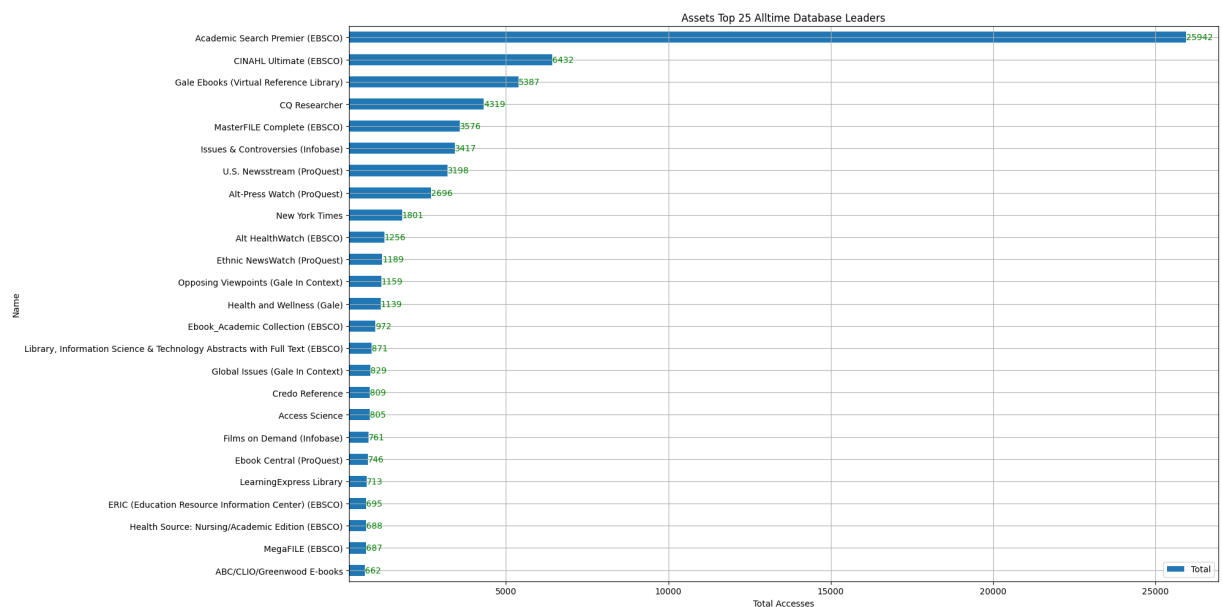
Enter number wanted: 25

Out[12]:

	ID	Name	Type	Total
0	3361064	ABC/CLIO/Greenwood E-books	Database	662
1	3361125	MegaFILE (EBSCO)	Database	687
2	3361112	Health Source: Nursing/Academic Edition (EBSCO)	Database	688
3	3361098	ERIC (Education Resource Information Center) (...	Database	695
4	10410784	LearningExpress Library	Database	713
5	3361095	Ebook Central (ProQuest)	Database	746
6	3361102	Films on Demand (Infobase)	Database	761
7	3361066	Access Science	Database	805
8	3361093	Credo Reference	Database	809
9	3361106	Global Issues (Gale In Context)	Database	829
10	3361121	Library, Information Science & Technology Abst...	Database	871
11	3361163	Ebook_Academic Collection (EBSCO)	Database	972
12	3361109	Health and Wellness (Gale)	Database	1139
13	3361135	Opposing Viewpoints (Gale In Context)	Database	1159
14	3361099	Ethnic NewsWatch (ProQuest)	Database	1189
15	3361068	Alt HealthWatch (EBSCO)	Database	1256
16	22812778	New York Times	Database	1801
17	3361071	Alt-Press Watch (ProQuest)	Database	2696
18	3361145	U.S. Newsstream (ProQuest)	Database	3198
19	3361115	Issues & Controversies (Infobase)	Database	3417
20	3361124	MasterFILE Complete (EBSCO)	Database	3576
21	3361091	CQ Researcher	Database	4319
22	3361103	Gale Ebooks (Virtual Reference Library)	Database	5387
23	3361083	CINAHL Ultimate (EBSCO)	Database	6432
24	3361065	Academic Search Premier (EBSCO)	Database	25942

In [13]:

```python
ax = database_leaders.plot(
        x="Name",
        y="Total",
        figsize=(20,10),
        kind="barh",
        grid=True,
        xlim=[database_leaders["Total"].min()-500, database_leaders["Total"].max()+1000],
        xlabel="Total Accesses",
        title=f"Assets Top {num_leaders} Alltime Database Leaders")

ax.bar_label(ax.containers[0], color="green")
plt.tight_layout()

plt.savefig(f"Assets/db_assets_{num_leaders}.png")

plt.show()
```


In [14]:

```python
print(f'Mean of Asset Views: {assets_df["Total"].mean()}')
print(f'Median of Asset Views: {assets_df["Total"].median()}')
print(f'Mode of Asset Views: {st.mode(assets_df["Total"], axis=None, keepdims=False)}')

print(f'Min of Asset Views: {assets_df["Total"].min()}')
print(f'Max of Asset Views: {assets_df["Total"].max()}')
```

```
Mean of Asset Views: 90.2726480836237
Median of Asset Views: 3.0
Mode of Asset Views: ModeResult(mode=1, count=349)
Min of Asset Views: 1
Max of Asset Views: 25942
```