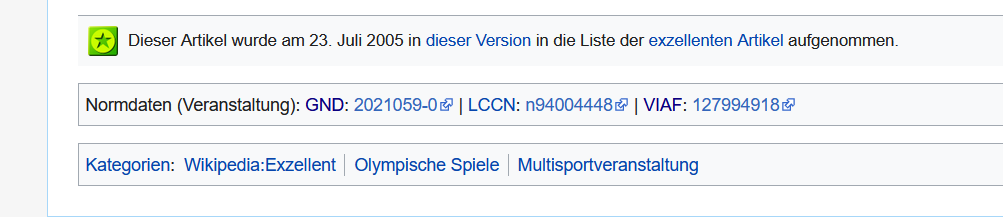
**Wir wollen alle Athleten (Name, Sportart, Alter, Herkunftsland) von Olympische Spielen, die im 20. Jahrhundert in Europa stattgefunden haben. Wir wollen dazu den Query Service von Wikidata nutzen.**

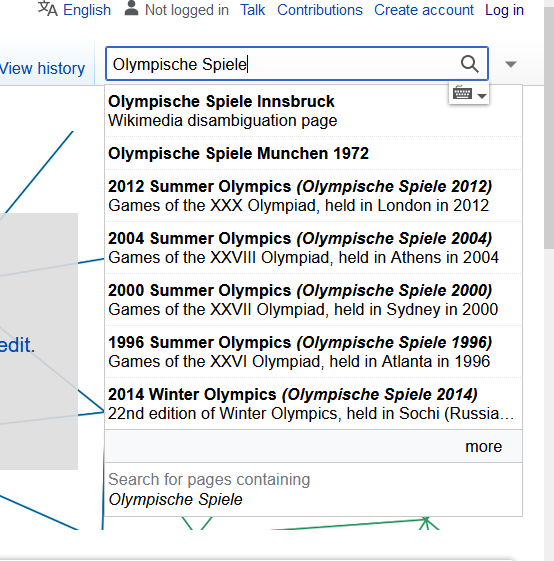
Beginnen können wir etwa damit, dass wir uns alle Olympischen Spiele einmal holen. Dafür brauchen wir einmal die Q-Nummer für Olympischen Spiele

IM Wikipediaeintrag (<https://de.wikipedia.org/wiki/Olympische_Spiele>) finden wir unten oft Normdaten-IDs.

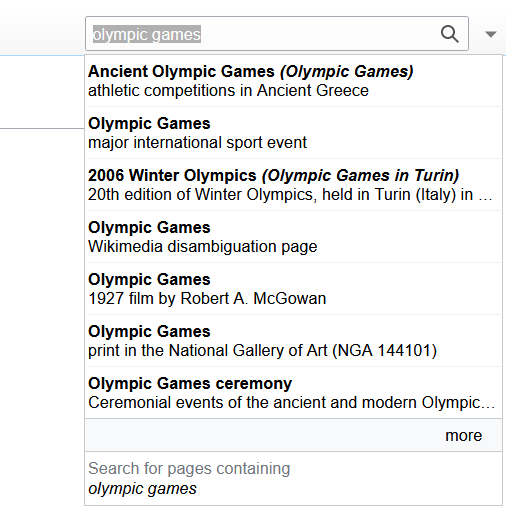


Hier ist gerne auch eine Wikidata Q-Nummer dabei.

Diesmal leider nicht. Dann verwende ich gerne die Suche (Autocomplete) rechts oben) in Wikidata



Auch das hat nicht direkt funktioniert. Vielleicht ist es besser den englischen Begriff “olympic games” zu verwenden:



Und der zweite Treffer schaut vielversprechend aus.

Also dieses Konzept: <https://www.wikidata.org/wiki/Q5389>

Im Query Srvice kann man mit STRG+Leertaste auch dieses Autocomplete aufrufen.

Als nächsten Schritt wollen wir alle verknüpften Daten anschauen, die mit Q5389 in Verbindung stehen.

Mit dieser Query bekomme ich alle Properties, die auf Literale oder Q-Nummern verweisen, die von wd:Q5389 “weggehen”.

SELECT \* WHERE {

wd:Q5389 ?p ?s.

}

Mit dieser Query bekomme ich alle Q-Nummern, die über welche Proeprties auch immer auf wd:Q5389 zeigen.

SELECT \* WHERE {

?o ?p wd:Q5389.

}

Um jetzt alle Olympischen Spiele in Europa im 20.Jahrhundert zu bekommen (entschuldigt den Euro-Zentrismus) müssen wir die Properties finden, die den Ort und die Zeit der Spiele definieren. Sehr hilfreich sind, dann immer Beispiele, also Q-Nummern in Wikidata.

Zum Beispiel <https://www.wikidata.org/wiki/Q9672> . Dort finden wir [point in time](https://www.wikidata.org/wiki/Property:P585) (P585), [start time](https://www.wikidata.org/wiki/Property:P580) (P580) die geeignte Proeprties für das Datum sein können und [country](https://www.wikidata.org/wiki/Property:P17) (P17) bzw. location (P276), um den Ort zu finden.

Mit dieser Query würden wir alle Olympischen Spiele mit Ort und Datum bekommen.

SELECT \* WHERE {

?o ?p wd:Q5389;

wdt:P17 ?country.

wdt:P580 ?time.

}

Es sind leider nur zwei Treffer. Vielleicht passt unser wd:Q5389 nicht?

Wenn wir uns <https://www.wikidata.org/wiki/Q8438>, dann ist es eine [instance of](https://www.wikidata.org/wiki/Property:P31) (P31) <https://www.wikidata.org/wiki/Q159821> (Summer Olympic Games) und das ist eine [subclass of](https://www.wikidata.org/wiki/Property:P279) (P279) <https://www.wikidata.org/wiki/Q5389>.

Also wd:Q5389 passt schon, wir müssen aber nur auch die Unterklassen davon berücksichtigen.

<https://www.wikidata.org/wiki/Q82414> sind die Winter Olympic Games

Holen wir uns alle Unterklassen von wd:Q5389 und geben sie in die Variabel ?olympic\_game\_classes.

SELECT \* WHERE {

?olympic\_games wdt:P279 wd:Q5389.

}  
  
Jetzt wollen wir alle Instanzen von all diesen Klassen:

SELECT ?olympic\_games WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

}

Und schaut gut aus. Beim Reinschaun in ein paar Q-Nummern sind es immer olympische Spiele. Jetzt wollen wir Datum und Ort auch mit abfragen. Vorsichtshalber aber in einem OPTIONAL, damit wir auch die Einträge finden, bei denen Ort oder Datum bzw. Ort und Datum nicht angegeben sind.

SELECT ?olympic\_games ?country ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

OPTIONAL{?olympic\_games wdt:P17 ?country}

OPTIONAL{?olympic\_games wdt:P580 ?time}

}

Bei vielen sehen, wir, dass P:17 bzw. P:580 nicht angegeben sind. Das sind dann so Einträge wie <https://www.wikidata.org/wiki/Q57337450>, also die Spiele im Jahre 772 v. Chr.

Wir haben ja auch gesehen, dass es noch location (P276) und point in time (P585) als Proprties gibt. Wir können also ein OR für Properties verwenden und auch diese Felder mit der anderen Dataproperty befüllen.

SELECT ?olympic\_games ?country ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

OPTIONAL{?olympic\_games wdt:P17|P276 ?country}

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

}

Diese Treffermegne wollen wir jetzt auf “Europa” und auf 20.Jahrhundert eingrenzen.

Wir brauchen einmal die Jahre. Mit year(xsd:dateTime(?time)) as ?year) kann man aus dem Datentyp dateTime mit der Funkltion year() die jahreszahl rausholen und das Ergebnis in ?year binden. Wir zeigen uns das einfach mal an.

SELECT ?olympic\_games ?country (year(xsd:dateTime(?time)) as ?year) ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

OPTIONAL{?olympic\_games wdt:P17|wdt:P276 ?country}

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

}

Mit dieser Query haben wir dann alle Olympischen Spiele , die zwischen 1899 und 2000 statt gefunden haben:

SELECT ?olympic\_games ?country ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

OPTIONAL{?olympic\_games wdt:P17|wdt:P276 ?country}

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

}

In ?country haben wir jetzt treffer wie <https://www.wikidata.org/wiki/Q20> und <https://www.wikidata.org/wiki/Q866409> für das Bislett Stadion. Vl ist es hier doch besser nur die country Property zu verwenden, weil damit eben Norwegen gefunden wird und mit location eben auch das Stadion. Aber genau wissen wir es gerade nicht. Deswegen holen wir uns in eine neue Variabel alle Q-Nummern, die [instance of](https://www.wikidata.org/wiki/Property:P31) https://www.wikidata.org/wiki/Q6256 und die [part of](https://www.wikidata.org/wiki/Property:P361) [Europe](https://www.wikidata.org/wiki/Q46) sind.

Und, wenn wir hier die gleiche Variable wie im OPTIONAL nutzen und durch die konjunktive Logik von SPARQL haben wir auch schon das Ergebnis.

SELECT ?olympic\_games ?country ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?country wdt:P31 wd:Q6256;

wdt:P361 wd:Q46.

OPTIONAL{?olympic\_games wdt:P17|wdt:P276 ?country}

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

}

Jetzt schauen wir weiter, was wir noch über mit den gefunden Results machen können und zeigen uns alle Properties auf alle Q-Nummern|Literale an.

SELECT ?olympic\_games ?a ?b

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?country wdt:P31 wd:Q6256;

wdt:P361 wd:Q46.

OPTIONAL{?olympic\_games wdt:P17|wdt:P276 ?country}

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

?olympic\_games ?a ?b.

}

Wir werden rdfs:label nutzen, um statt den Q-Nummern was Lesbares für uns zu haben. Haben wir aber lieber in einem OPTIONAL. Und wir holen uns das rdfs:label mit dem “en” language tag. Wir machen das für ?olympic\_games und auch für ?country.

SELECT ?olympic\_games\_label ?country\_label ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?country wdt:P31 wd:Q6256;

wdt:P361 wd:Q46.

OPTIONAL{?olympic\_games wdt:P17|wdt:P276 ?country}

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

OPTIONAL{?olympic\_games rdfs:label ?olympic\_games\_label}

FILTER ( lang(?olympic\_games\_label) = "en" )

OPTIONAL{?country rdfs:label ?country\_label}

FILTER ( lang(?country\_label) = "en" )

}

Jetzt haben wir aber das Problem, dass wir nur Italy, Switzerland und Norway in der “country\_label”-Spalten haben. Das passt nicht!

Das Problem ist, dass mit “?country wdt:P31 wd:Q6256; wdt:P361 wd:Q46.” nicht alle europäischen Länder selektiert werden. Z.B. hat [Soviet Union](https://www.wikidata.org/wiki/Q15180) hier keine Verbindung. Wir müssen also irgendwie allgemeiner werden.

Wir müssen da wirklich sehr allgemein sein. Und nehmen alles, das eine “?country ?someConncetionTo wd:Q46." mit Europa hat.

SELECT ?olympic\_games ?olympic\_games\_label ?country ?country\_label ?time

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

OPTIONAL{?country rdfs:label ?country\_label}

FILTER ( lang(?country\_label) = "en" )

OPTIONAL{?olympic\_games rdfs:label ?olympic\_games\_label}

FILTER ( lang(?olympic\_games\_label) = "en" )

}

Jetzt lassen wir uns die Teilnehmeranzahl und die TeilnehmerInnen anzeigen, die vom Olympischen Spiel aus damit verknüpft sind (es könnten aber Athleten eben auch mit dem Spiel verknüpft sein. Stichwort: gerichteter Graph!)

Damit bekommen wir aber auch historische Staaten, wie etwa <https://www.wikidata.org/wiki/Q209065>. Sowas muss man eben immer bedenken. Aber es schaut sonst ganz gut aus.

SELECT ?olympic\_games ?olympic\_games\_label ?country ?country\_label ?time ?number\_of\_participants ?participants

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

OPTIONAL{?country rdfs:label ?country\_label}

FILTER ( lang(?country\_label) = "en" )

OPTIONAL{?olympic\_games rdfs:label ?olympic\_games\_label}

FILTER ( lang(?olympic\_games\_label) = "en" )

OPTIONAL{?olympic\_games wdt:P1132 ?number\_of\_participants}

OPTIONAL{?olympic\_games wdt:P710 ?participants}

}

Mit [participant in](https://www.wikidata.org/wiki/Property:P1344) [1952 Winter Olympics](https://www.wikidata.org/wiki/Q9623) wird nämlich vom Athleten aus auf ein Olympisches Spiel referenziert. **Folgende QUERY hängt sich auf!!!** Wenn wir hier die andere Richtung der TeilnehmerInnen mitnehmen. Das geht nämlich ordentlich in die Performance, weil es einfach viele Athleten sind.

SELECT ?olympic\_games\_label ?country\_label ?time ?number\_of\_participants ?participants ?participants\_other\_direction

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

OPTIONAL{?country rdfs:label ?country\_label}

FILTER ( lang(?country\_label) = "en" )

OPTIONAL{?olympic\_games rdfs:label ?olympic\_games\_label}

FILTER ( lang(?olympic\_games\_label) = "en" )

OPTIONAL{?olympic\_games wdt:P1132 ?number\_of\_participants}

OPTIONAL{?olympic\_games wdt:P710 ?participants}

OPTIONAL{?participants\_other\_direction wdt:P1344 ?olympic\_games}

}

Unser Ziel war ja:

*Wir wollen alle Athleten (Name, Sportart, Alter, Herkunftsland) von Olympische Spielen, die im 20. Jahrhundert in Europa stattgefunden haben. Wir wollen dazu den Query Service von Wikidata nutzen.*

**So bekommen wir alle Athleten, die an Olympischen Spielen teilgenommen haben, die in Europa und im 20. Jahrhundert stattfanden.**

SELECT ?participants

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

{?olympic\_games wdt:P710 ?participants.}

UNION

{?participants wdt:P1344 ?olympic\_games}

}

“participant in” bzw. “participats” muss aber auch nicht heißen, dass es SportlerInnen sind. Könnten auch OrganisatorInnen sein. Wir können also eventuell mit [occupation](https://www.wikidata.org/wiki/Property:P106)

[athlete](https://www.wikidata.org/wiki/Q2066131) weiter einschränken. **Und zur Disziplin kommen wir mit** [**sport**](https://www.wikidata.org/wiki/Property:P641)**.**

Es gibt dann ganz viele Q-Nummern für Occupation, nämlich <https://www.wikidata.org/wiki/Q10843402>, <https://www.wikidata.org/wiki/Q45199>, <https://www.wikidata.org/wiki/Q2066131>, etc.

Wir werden aber einfach [sport](https://www.wikidata.org/wiki/Property:P641) verwenden. Es gibt auch einen eigenen **SERVICE von Wikidata, der uns die rdfs:label gleich in englischer Sprache anzeigt**.

Wir setzen aber jetzt ein LIMIT, weil wir so viele Athleten haben.

SELECT DISTINCT ?participantLabel ?sportLabel

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

{?olympic\_games wdt:P710 ?participant.}

UNION

{?participant wdt:P1344 ?olympic\_games}

?participant wdt:P641 ?sport.

SERVICE wikibase:label {

bd:serviceParam wikibase:language "en" .

}

}

LIMIT 5000

Mit <https://www.wikidata.org/wiki/Property:P569> und unserer “time”-Variabel können wir nun das **Alter zur Zeit der Olympiade ausrechnen**, wenn es die Daten gibt. Dafür können wir ein neues Ergebnis mit BIND in eine neue Variabel schreiben. Geht wieder ohne LIMIT, weil wir viel weniger Treffer haben. Ist aber jetzt natürlich nicht. Jetzt gibt es in den Daten Fälle wie birth = 19. Jahrhundert. Hier bekommen wir jetzt eine 0 in der “age”-Spalte.

SELECT DISTINCT ?participantLabel ?sportLabel ?age

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

{?olympic\_games wdt:P710 ?participant.}

UNION

{?participant wdt:P1344 ?olympic\_games}

?participant wdt:P641 ?sport;

wdt:P569 ?birth;

wdt:P570 ?death.

BIND( year(xsd:dateTime(?death))-year(xsd:dateTime(?birth)) as ?age )

SERVICE wikibase:label {

bd:serviceParam wikibase:language "en" .

}

}

**Und mit wdt:P27 haben wir dei Staatsangehörigkeit auch noch.**

SELECT DISTINCT ?participantLabel ?sportLabel ?citizenshipLabel ?age

WHERE {

?olympic\_game\_classes wdt:P279 wd:Q5389.

?olympic\_games wdt:P31 ?olympic\_game\_classes.

?olympic\_games wdt:P17 ?country.

?country ?someConncetionTo wd:Q46.

OPTIONAL{?olympic\_games wdt:P580|wdt:P585 ?time}

FILTER(year(xsd:dateTime(?time)) > 1899 && year(xsd:dateTime(?time)) < 2000 )

{?olympic\_games wdt:P710 ?participant.}

UNION

{?participant wdt:P1344 ?olympic\_games}

?participant wdt:P641 ?sport;

wdt:P569 ?birth;

wdt:P570 ?death.

OPTIONAL{?participant wdt:P27 ?citizenship.}

BIND( year(xsd:dateTime(?time))-year(xsd:dateTime(?birth)) as ?age )

SERVICE wikibase:label {

bd:serviceParam wikibase:language "en" .

}

}