Druid & Rune Cheat Sheet

by tmp_7890 via cheatography.com/123074/cs/23077/

How to connect to Druid?

Port-forward to the Druid cluster

```
kubectl port-forward -n
wiremind-druid-{env} druid-
{env}-broker-{id-container}
8082:8082
```

Exemple: `kubectl port-f orward -n wiremi nd- dru id- staging druid - sta gin g-b rok er- 795 c75 64 5 6-sjbm9 8082:8082`

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

Optional: configure a SQL client

If you want to query Druid with an SQL client, you can do so using this driver.

Optional: configure a SQL client (cont)

In Database URL, specify: `jdbc:a vat ica :re mot e:u rl= htt p:/ /lo cal - hos t:8 082 /dr uid /v2 /sq l/a vatica/`

Tools: DBVizualizer, DBeaver, DataGrip, ...

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

What data sources can I query?

fact_passenger_event

Granularity	od_id **X** event_type **X** ticket_key

fact_passenger_event (cont)

Refresh	daily

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.
Page 3 of 100.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

fact_passenger_event (cont)

Time reference (*__time*)	Exact timestamp of the event.

event_type can be: `abando n_cart`, `conf ir m_p urchase`, `cancel _pu rchase`

fact_daily_od_bucket

Granularity	od_id **X** bucket_id **X** day_x

fact_daily_od_bucket (cont)

Refresh	daily

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.
Page 4 of 100.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

fact_daily_od_bucket (cont)	
Time reference (*__time*)	UTC time at train timezone midnight corresponding to this *day_x*

fact_daily_leg_physical_inventory	
Granularity	leg_id **X** physical_inventory_id **X** day_x

fact_daily_leg_physical_inventory (cont)	
Refresh	daily

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.
Page 5 of 100.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

fact_daily_leg_physical_inventory (cont)

Time reference (*__time*)	UTC time at train timezone midnight corresponding to this *day_x*

Query tips

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

Useful SQL functions

Cast datetime to date	CAST(departure_datetime TO DATE)

Most common columns

__time	If possible, **filter on**.

Most common columns (cont)

day_x	Compute using the sevice timezone.

By **tmp_7890**

cheatography.com/tmp-7890/

Not published yet.

Last updated 5th June, 2020.

Page 7 of 100.

Sponsored by **ApolloPad.com**

Everyone has a novel in them. Finish Yours!

https://apollopad.com

Most common columns (cont)

[origin/destination]_station_[id/name/zone]

Most common columns (cont)

[departure/arrival]_datetime

Most common columns (cont)

service_[id/number/status]

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

Most common columns (cont)

market_[id/name]

Most common columns (cont)

ordered_leg_ids	You must use ARRAY functions.

Most common columns (cont)

od_[capacity/lid/id]

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

Most common columns (cont)

bucket_name

Most common columns (cont)

availability[_physical]_seats_[start/end]_day

Most common columns (cont)

[cumulative_]sum_[cancelled/confirmed/net]_bookings

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

Most common columns (cont)

[cumulative_]sum_[net]_revenue_vat_[-inc/exc]	

Most common columns (cont)

has_event-_occurred	If the line is exactly the same as the day before.

Most common columns (cont)

price_-vat_[i-nc/exc]	For fact_passenger_event, it might be 0 (abandon_cart) or negative (cancelled_purchase). Use base_price_vat_inc for initial price.

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

Most common columns (cont)

travel_time_minute

More information here.

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

By **tmp_7890**
cheatography.com/tmp-7890/

Not published yet.
Last updated 5th June, 2020.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com