Queries for Slide 1-

Total orders:

```
SELECT sum(orders)
FROM `dh-interview-case-study.logistics_service_product_analytics.orders`
```

Total delays:

```
SELECT count(distinct event_id)
FROM `dh-interview-case-study.logistics_service_product_analytics.events`
```

Total vouchers:

```
SELECT count(distinct voucher_id)
FROM `dh-interview-case-study.logistics_service_product_analytics.events`
```

Data considerations for computing total survey responses:

Refer to the 'DeliveryHero_CaseStudy.ipynb'- ignored flags with multiple in either survey flag or event flag as per the assumption.

Total survey responses:

computed on excel 'country split' excel file, base for that data-

```
With basetable1 as
(
SELECT  a.customer_id as customer,*
 FROM `dh-interview-case-study.logistics_service_product_analytics.customer_suveys` as a
inner join  `dh-interview-case-study.logistics_service_product_analytics.events` as b
on a.customer_id = b.customer_id and survey_response_ts > event_ts
where date(survey_response_ts) between date('2020-07-01') and date('2020-07-31') ),

basetable2 as (
SELECT  customer,count(survey_id)
 from basetable1
group by 1
having count(survey_id) =1),

basetable3 as (
```

```
select basetable2.customer,basetable1.survey_id
from basetable2
inner join basetable1
on basetable2.customer = basetable1.customer)
```

```
select c.customer,action_type,country,event_type,c.survey_id,survey_rating
from basetable1 as c
inner join basetable3 as d
on c.customer = d.customer and c.survey_id = d.survey_id
where DATE_DIFF(date(survey_response_ts), date(event_ts), DAY) between 0 and 5
```

Total satisfactory responses:

Computed through the country split excel file

For Slide 2-

For delay rate:

used the below query and divided by total orders calculated by similar logic in orders table-

```
SELECT  EXTRACT(WEEK FROM  event_ts AT TIME ZONE "UTC") AS
the_week_utc,count(distinct event_id)
FROM `dh-interview-case-study.logistics_service_product_analytics.events`
group by 1
order by 1
```

For voucher provision rate:

```
SELECT  EXTRACT(WEEK FROM  event_ts AT TIME ZONE "UTC") AS
the_week_utc,count(distinct voucher_id)/count(distinct event_id)
FROM `dh-interview-case-study.logistics_service_product_analytics.events`
group by 1
order by 1
```

For Slide 3-

Refer ‘country split’ excel file for calculations related to the table on this slide

Refer the ‘Excel file for Wow-cher’ file for calculations related to the hypothesis test

For Slide 4-

Refer to the ‘country split’ excel file

For Slide 5-

Refer to the ‘customer value index split’ excel file, base for the data-

```
With basetable1 as
(
SELECT  a.customer_id as customer,*
 FROM `dh-interview-case-study.logistics_service_product_analytics.customer_suveys` as a
inner join  `dh-interview-case-study.logistics_service_product_analytics.events` as b
on a.customer_id = b.customer_id and survey_response_ts > event_ts
where date(survey_response_ts) between date('2020-07-01') and date('2020-07-31') ),

basetable2 as (
SELECT  customer,count(survey_id)
 from basetable1
group by 1
having count(survey_id) =1),

basetable3 as (
select basetable2.customer,basetable1.survey_id
from basetable2
inner join basetable1
on basetable2.customer = basetable1.customer)


select c.customer,action_type,customer_value_index,event_type,c.survey_id,survey_rating
from basetable1 as c
inner join basetable3 as d
on c.customer = d.customer and c.survey_id = d.survey_id
where DATE_DIFF(date(survey_response_ts), date(event_ts), DAY) between 0 and 5
and customer_value_index != -1
```