Here is my report :

https://app.mode.com/datafra75/reports/fbbbbb90fd57

The SQL can be viewed at : (on the left)

https://app.mode.com/datafra75/reports/fbbbbb90fd57/details

The interpretation of the p-value can be found in the SQL comments of the fifth question, or below.

In cas there is a problem, here are the results... I hope it's readable. Thank you.

Question number 1

```
--We are running an experiment at an item-level, which means all users who visit will see the
same page, but the layout of different item pages may differ.
--Compare this table to the assignment events we captured for user_level_testing.
--Does this table have everything you need to compute metrics like 30-day view-binary?

SELECT
  *
FROM
  dsv1069.final_assignments_qa

-- No, assignment dates and times are missing
```

Question number 2

```
--Reformat the final_assignments_qa to look like the final_assignments table, filling in any
missing values with a placeholder of the appropriate data type.

-- There are no NULL values in any column of final_assignments_qa and all the values are 0 or 1
-- It lacks the date so I add it
SELECT item_id,
        test_a
AS test_assignment,
        'test_a'
AS test_number,
        CAST('2021-01-01 00:00:00' AS timestamp)         AS test_date
FROM dsv1069.final_assignments_qa
UNION
SELECT item_id,
        test_b
AS test_assignment,
        'test_b'
AS test_number,
        CAST('2021-01-01 00:00:00' AS timestamp)         AS test_date
```

```
FROM dsv1069.final_assignments_qa
UNION
SELECT item_id,
        test_c
AS test_assignment,
        'test_c'
AS test_number,
        CAST('2021-01-01 00:00:00' AS timestamp)         AS test_date
FROM dsv1069.final_assignments_qa
UNION
SELECT item_id,
        test_d
AS test_assignment,
        'test_d'
AS test_number,
        CAST('2021-01-01 00:00:00' AS timestamp)         AS test_date
FROM dsv1069.final_assignments_qa
UNION
SELECT item_id,
        test_e
AS test_assignment,
        'test_e'
AS test_number,
        CAST('2021-01-01 00:00:00' AS timestamp)         AS test_date
FROM dsv1069.final_assignments_qa
UNION
SELECT item_id,
        test_f
AS test_assignment,
        'test_f'
AS test_number,
        CAST('2021-01-01 00:00:00' AS timestamp)         AS test_date
FROM dsv1069.final_assignments_qa
```

Question number 3

```
-- Use this table to
-- compute order_binary for the 30 day window after the test_start_date
-- for the test named item_test_2

SELECT
  test_assignment,
  COUNT(DISTINCT item_id) AS items_count,
  SUM(order_binary) AS order_binary_30d_sum
FROM
```

```
(
    SELECT
      final_assignments.item_id,
      final_assignments.test_assignment,
      final_assignments.test_start_date,
      orders.created_at,
      MAX(CASE
        WHEN (orders.created_at > final_assignments.test_start_date
        AND DATE_PART('day', orders.created_at - final_assignments.test_start_date) <= 30) -- (>=0 because You may include the day the test started")
        THEN 1
        ELSE 0
      END) AS order_binary
    FROM dsv1069.final_assignments AS final_assignments
    LEFT OUTER JOIN dsv1069.orders AS orders
    ON final_assignments.item_id = orders.item_id
    WHERE test_number = 'item_test_2'
    GROUP BY
      final_assignments.item_id,
      final_assignments.test_assignment,
      final_assignments.test_start_date,
      orders.created_at
) AS order_binary_table
GROUP BY
  test_assignment
```

Question number 4

```
-- Use this table to
-- compute view_binary for the 30 day window after the test_start_date
-- for the test named item_test_2

SELECT
    test_assignment,
    COUNT(DISTINCT item_id) AS views_count,
    SUM(view_binary_30d) AS view_binary_30d_sum,
    AVG(view_binary_30d) AS view_binary_30d_avg
FROM
  (
  SELECT
      assignments.item_id,
      assignments.test_assignment,
      MAX(CASE
            WHEN (views.event_time > assignments.test_start_date
                   AND DATE_PART('day', views.event_time - assignments.test_start_date) <= 30)
-- (>=0 because You may include the day the test started")
            THEN 1
            ELSE 0
            END) AS view_binary_30d
  FROM dsv1069.final_assignments AS assignments
  LEFT JOIN dsv1069.view_item_events AS views
  ON assignments.item_id=views.item_id
  WHERE assignments.test_number='item_test_2'
  GROUP BY
     assignments.item_id,
     assignments.test_assignment
  ORDER BY
     item_id
  ) AS view_binaire_table
GROUP BY test_assignment
```

Question number 5

```
--Use the https://thumbtack.github.io/abba/demo/abba.html to compute the lifts in metrics and the
p-values for the binary metrics (30 day order binary and 30 day view binary) using a interval 95%
confidence.
```

AB testing for 30 day order binary:
-- control : 399/1130; treatment : 381/1068
-- lift : +1% (between -10% and +12%)
-- pval : 0.86
-- Interpretation : with a pval 0.86, the results are NOT statistically significant, we can't say if there was an increase of 1%...

AB testing for 30 day view binary:
-- control : 925/1130; treatment : 894/1068
-- lift : +2.3% (between -1.6% and +6.1%)
-- pval : 0.25
-- Interpretation : with a pval 0.25, the results are statistically significant, the treatment probably induced a lift of 2.3%