

Exercise Sheet 3

In this exercise sheet we're going to work on data preparation so that we have clean data to work with.

For these exercises we'll be using the Invoice table.

1 Exercise 1

1. Find the time difference each invoice and the start of 2009.

2 Exercise 2

1. Split up the InvoiceDate column into individual columns that contain the year, month, day, hour, minute, and second of each invoice.
2. Find all the events that happened within the first 5 days of each month.

3 Exercise 1 Sample Solutions

1.

```
SELECT "InvoiceId",
      "InvoiceDate",
      "InvoiceDate"::TIMESTAMP - '2009-01-01 00:00:00'::TIMESTAMP AS td
FROM "Invoice"
LIMIT 5;
```

4 Exercise 2 Sample Solutions

1.

```
SELECT "InvoiceId",
      "InvoiceDate",
      SUBSTR("InvoiceDate"::TEXT,1,4)::INT AS year,
      SUBSTR("InvoiceDate"::TEXT,6,2)::INT AS month,
      SUBSTR("InvoiceDate"::TEXT,9,2)::INT AS day,
      SUBSTR("InvoiceDate"::TEXT,12,2)::INT AS hour,
      SUBSTR("InvoiceDate"::TEXT,15,2)::INT AS minute,
      SUBSTR("InvoiceDate"::TEXT,18,2)::INT AS second
FROM "Invoice";
```
2.

```
SELECT "InvoiceId",
      "InvoiceDate",
      SUBSTR("InvoiceDate"::TEXT,1,4)::INT AS year,
      SUBSTR("InvoiceDate"::TEXT,6,2)::INT AS month,
      SUBSTR("InvoiceDate"::TEXT,9,2)::INT AS day,
      SUBSTR("InvoiceDate"::TEXT,12,2)::INT AS hour,
      SUBSTR("InvoiceDate"::TEXT,15,2)::INT AS minute,
      SUBSTR("InvoiceDate"::TEXT,18,2)::INT AS second
FROM "Invoice"
WHERE SUBSTR("InvoiceDate"::TEXT,9,2)::INT != 5;
```