

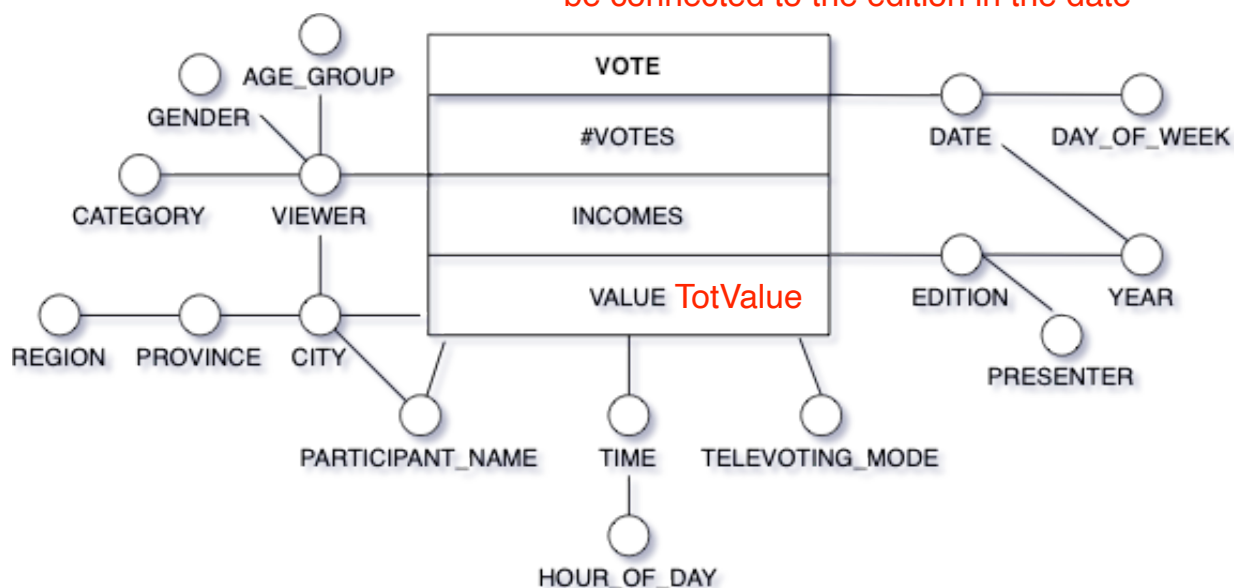
# Homework 1 - Enrico Castelli s280124

## Data Warehouse

---

### Conceptual schema

Year=Edition, so year is not needed and presenter can be connected to the edition in the date



### Logical schema

VIEWER(VID, AGEGROUP, GENDER, CATEGORY, LID)

LOCATION(LID, CITY, PROVINCE, REGION)

PARTICIPANT(PID, PARTICIPANTNAME, LID)

TIMETABLE(TID, TIME, HOUROFDAY)

EDITION(EID, PRESENTER, YEAR)

DATETABLE(DID, DATE, DAYOFWEEK, EID)

Facts table:

VOTE(VTID, VID, PID, TID, EID, DID, NUMVOTES, INCOMES, VALUE, TELEVOTINGMODE) **TelevotingMode should be a primary key**

## Queries

---

(a)

SQL

```

SELECT E.YEAR, VW.CITY, D.DAYOFWEEK,
       SUM(VT.INCOMES) OVER(ORDER BY E.YEAR
                           ROWS UNBOUNDED PRECEDING),
       100*SUM(VT.INCOMES)/SUM(SUM(VT.INCOMES)) OVER(PARTITION BY E.YEAR),
       AVG(SUM(VT.INCOMES)) OVER(PARTITION BY D.DAYOFWEEK)
FROM VOTE VT, VIEWER VW, DATETABLE D, EDITION E
WHERE VT.VTID=VW.VID AND VT.DID=D.DID AND D.EID=E.EID
      AND VT.TELEVOTINGMODE="Instagram"
GROUP BY E.YEAR, VW.CITY, D.DAYOFWEEK

```

(b)

SQL

```

SELECT E.YEAR, P.PARTICIPANTNAME,
       SUM(SUM(NUMVOTES)) OVER(PARTITION BY E.YEAR),
       100*SUM(NUMVOTES)/SUM(SUM(NUMVOTES)) OVER(PARTITION BY E.YEAR),
       RANK() OVER(PARTITION BY P.PARTICIPANTNAME
                   ORDER BY COUNT(DISTINCT D.DATE) DESC)
FROM VOTE VT, DATETABLE D, EDITION E, PARTICIPANT P
WHERE VT.DID=D.DID AND VT.PID=P.PID AND E.YEAR BETWEEN 2000 AND 2010
GROUP BY E.YEAR, P.PARTICIPANTNAME

```

(a)

```

SELECT
city,
SUM(SUM(income)) OVER (PARTITION BY city ORDER BY edition ROWS
UNBOUNDED PRECEDING),
SUM(income)/(SUM(SUM(income)) OVER (PARTITION BY edition),
SUM(income)/COUNT(DISTINCT date)
FROM
WHERE
GROUP BY edition, city

```

(b)

```

SELECT
edition,
participantName,
SUM(SUM(#votes)) OVER (PARTITION BY edition),
SUM(#votes)/SUM(SUM(#votes)) OVER (PARTITION BY edition),
RANK() OVER (PARTITION BY edition ORDER BY SUM(totVotes)/SUM(#votes)
DESC)
FROM
WHERE
GROUP BY participantName, edition

```