# DM & Visualization - Exam 2021-09-01 - Solution

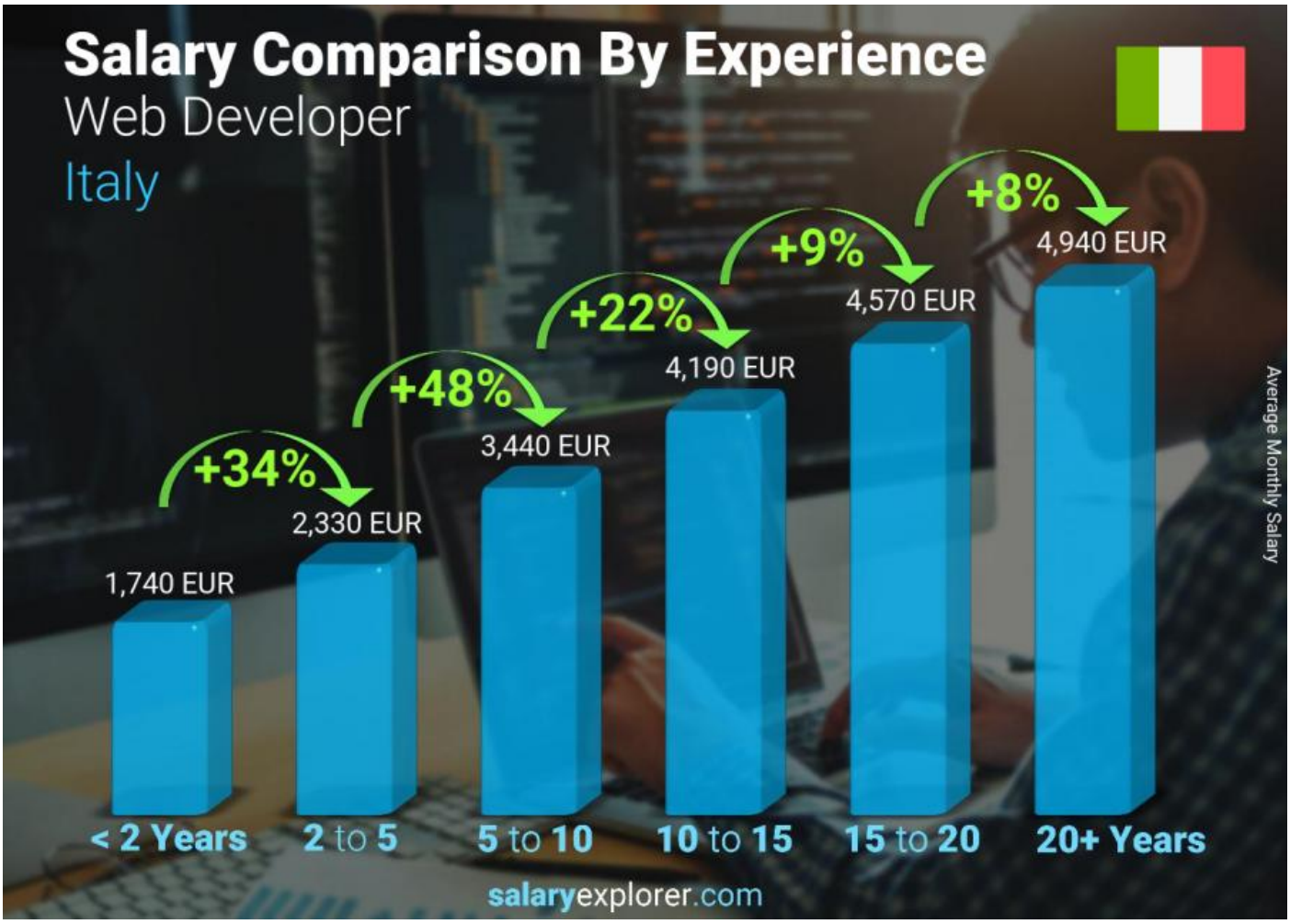



Figure 1: Salary Comparison By Experience

## Analysis

Analyze the above graph illustrating a comparison of different salaries by years of experience.

**Question: Is there one (or more) question addressed by the visualization?**

The question is quite clear: what is the average monthly salary of a web developer in Italy by years of experience?

**Data: Is the data quality appropriate?**

Accuracy: the values are monetary amounts and they are expressed in Euros.

Completeness: data are complete, as all the ranges of years of experience are available.

Consistency: data are not consistent, as the ranges have different timespans.

Currency: unclear, because no date is available.

Credibility: the source is Salary Explorer, data were probably collected with a survey.

Understandability: the values probably represent the salaries before taxation, but this is not reported.

Precision: precision is appropriate for the task.

**Visual Proportionality: Are the values encoded in a uniformly proportional way?**

The lengths of the bars are not proportional to the corresponding salaries because the y axis does not start from zero. In any case, the usage of 3D and the translucent effect make the comparison quite difficult.

**Visual Utility: All the elements in the graph convey useful information?**

Not at all: the background, the translucent effect, the shades, the percentages, the green arrows, the flag of Italy can be removed.

**Visual Clarity: Are the data in the graph clearly identifiable and understandable (properly described)?**

The numerical values associated with each bar are clearly reported. It is difficult to evaluate proportions and understand the correct value because of the 3D effect and the missing axis. It is stated that the values represent the average monthly salaries. The meaning of the percentages is clear, even if they are useless. Each salary is associated with a range of years of experience.

## Design

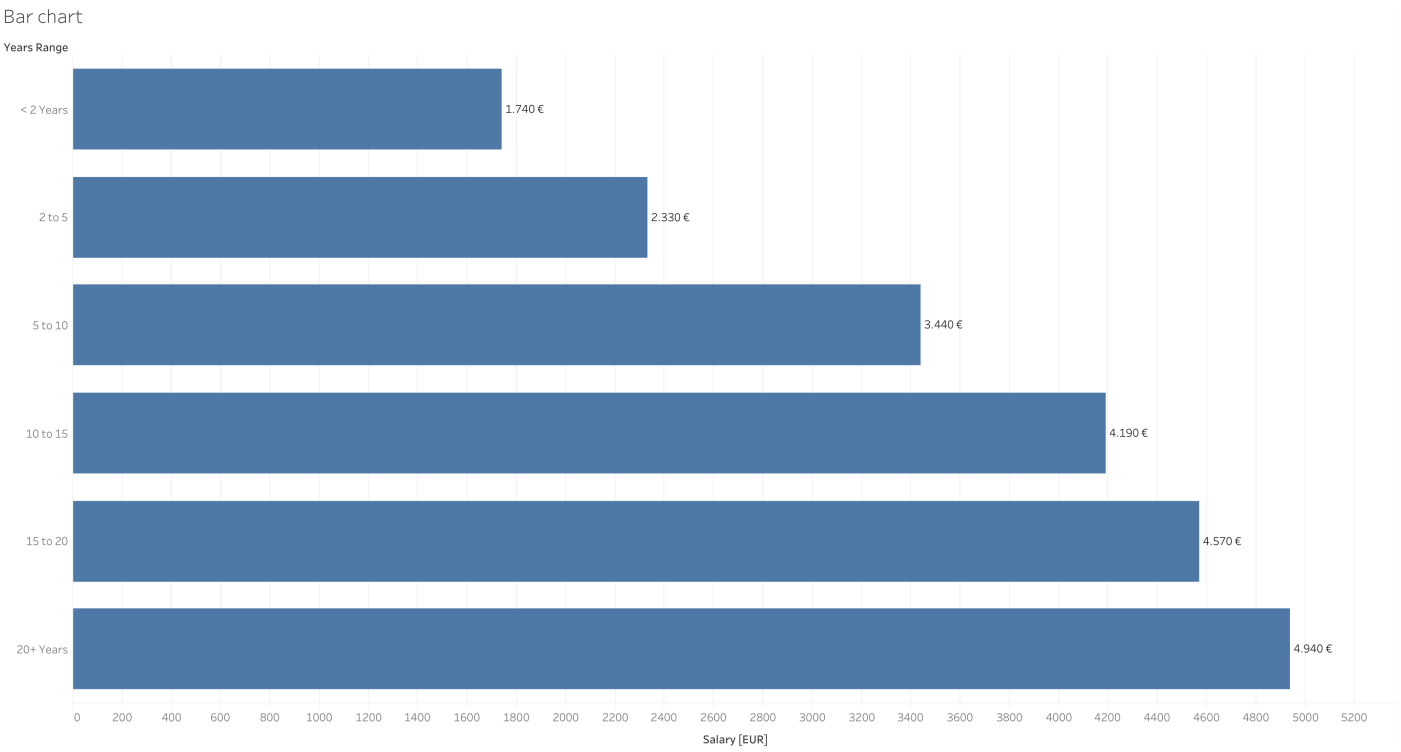
Design the visualization based on the following data structure

| Field | Dim./Measure | Description |
|---|---|---|
| YEARS_RANGE | Dimension | The years of experience |
| SALARY | Measure | The average monthly salary |

## Design schema

| Schema | Details |
|---|---|
| Columns: | SUM(SALARY) |
| Rows: | YEARS_RANGE |
| Graph type: | Bar |
| Color: | Default |
| Size: | Default |
| Label: | SUM(SALARY) |

## Sketch of the resulting graph

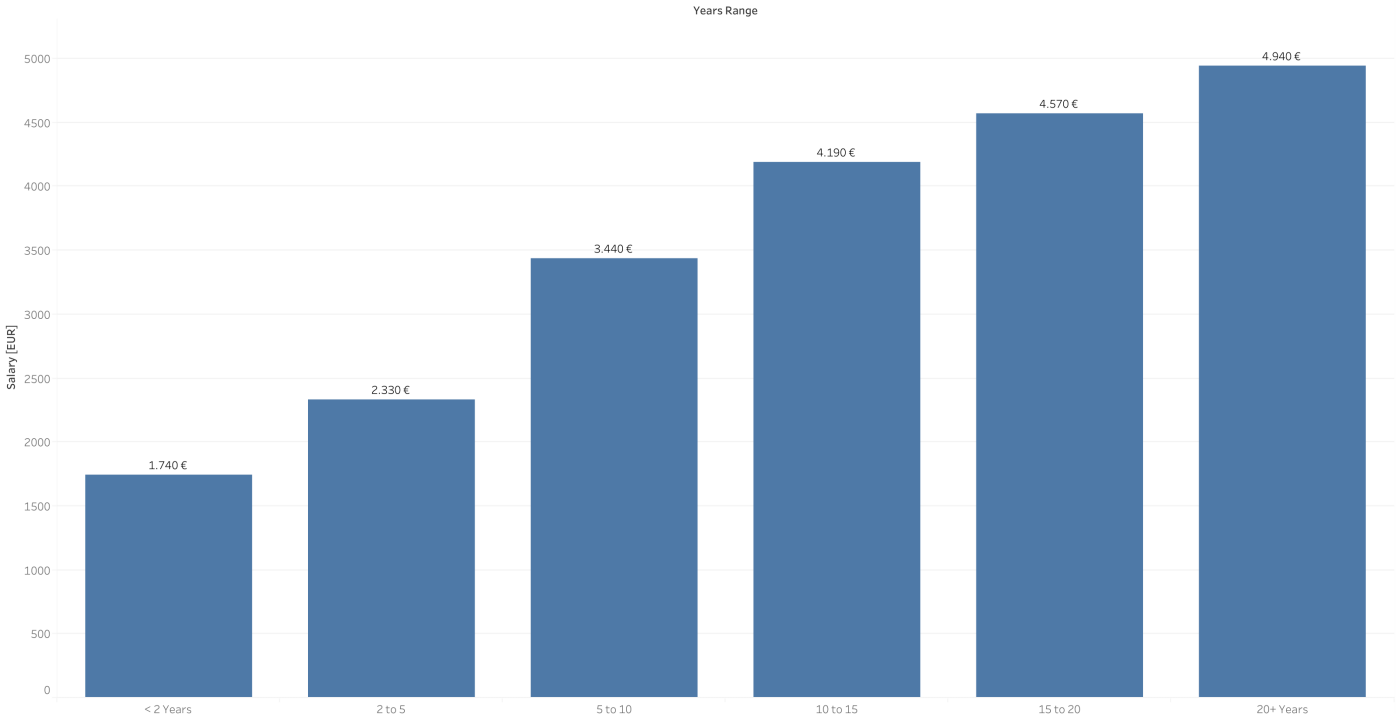


## Design schema

| Schema | Details |
|---|---|
| Columns: | YEARS_RANGE |
| Rows: | SUM(SALARY) |
| Graph type: | Bar |
| Color: | Default |
| Size: | Default |
| Label: | SUM(SALARY) |

## Sketch of the resulting graph

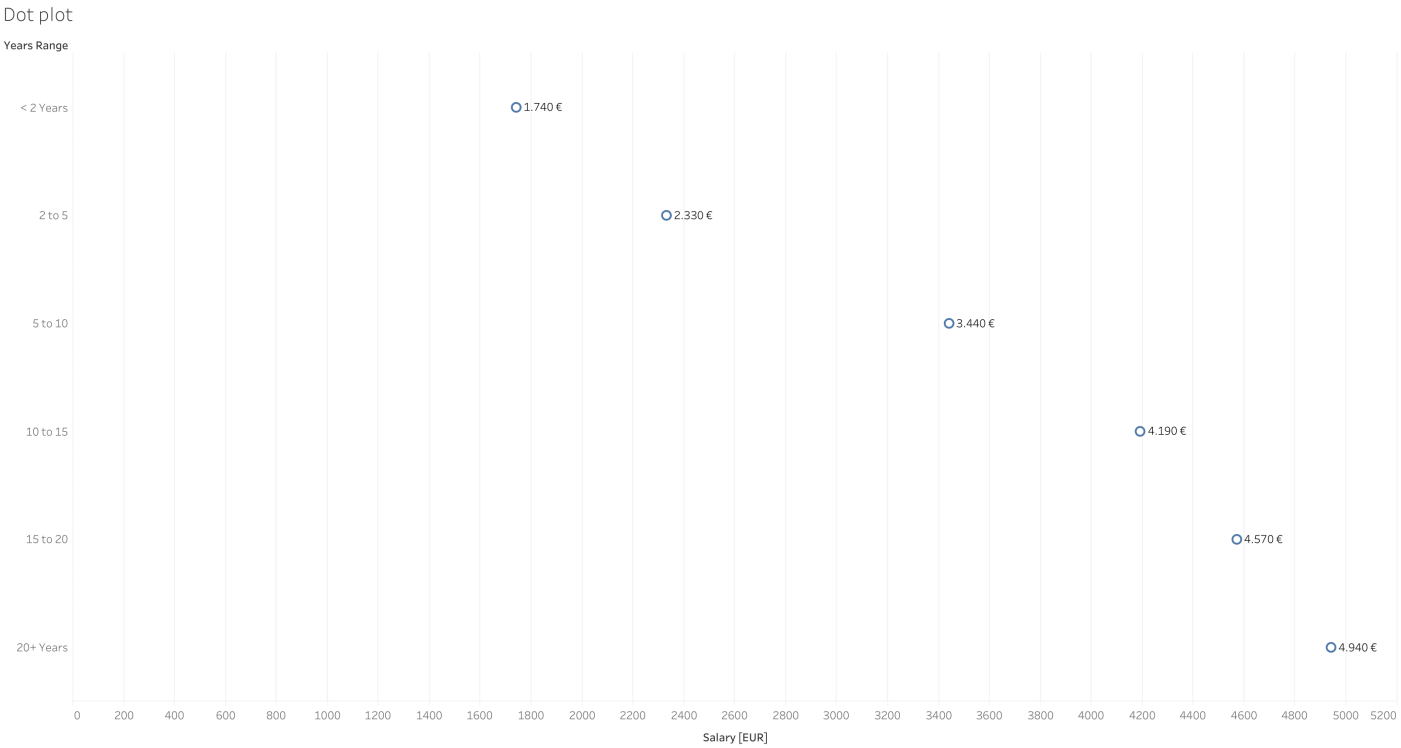
Vertical bar chart



## Design schema

| Schema | Details |
|---|---|
| Columns: | SUM(SALARY) |
| Rows: | YEARS_RANGE |
| Graph type: | Shape |
| Color: | Default |
| Size: | Default |
| Label: | SUM(SALARY) |

## Sketch of the resulting graph

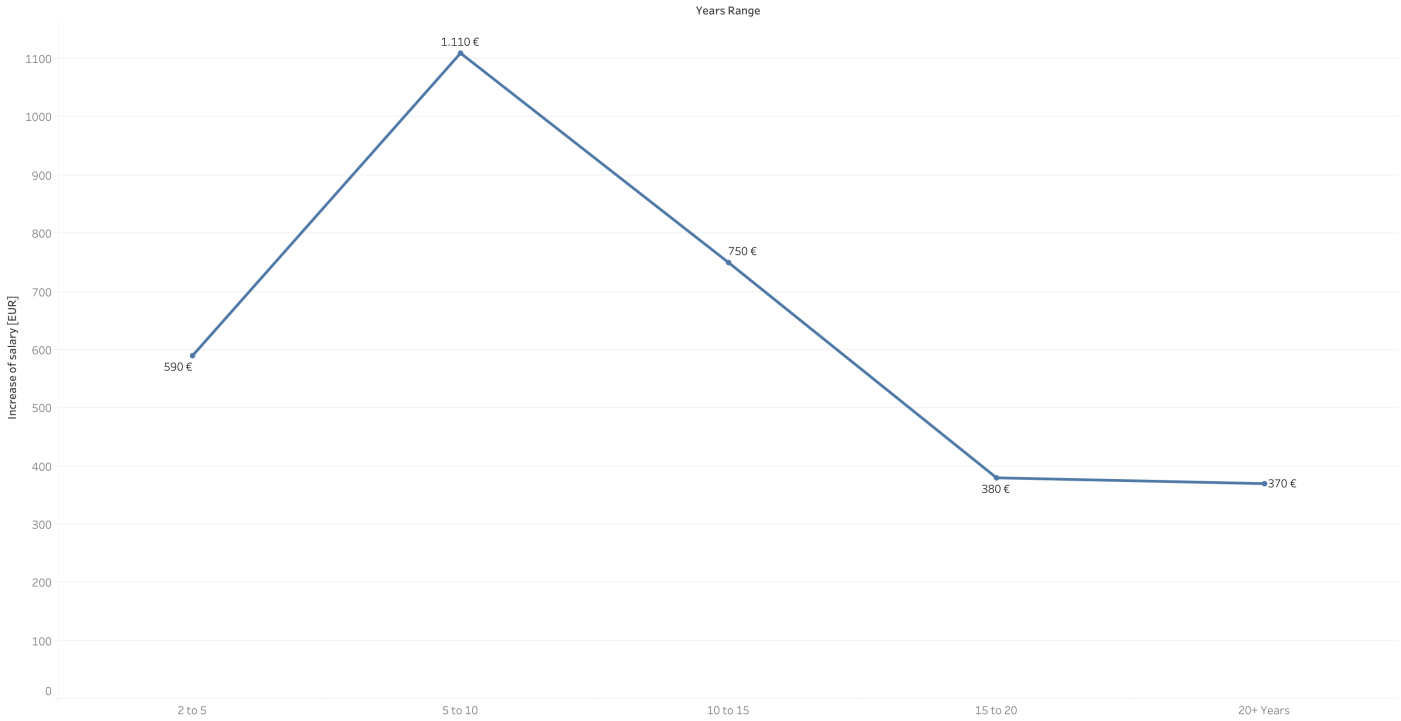


### Design schema

| Schema | Details |
|---|---|
| Columns: | YEARS_RANGE |
| Rows: | ZN(SUM(SALARY)) - LOOKUP(ZN(SUM(SALARY)), -1) |
| Graph type: | Line |
| Color: | Default |
| Size: | Default |
| Label: | ZN(SUM(SALARY)) - LOOKUP(ZN(SUM(SALARY)), -1) |

### Sketch of the resulting graph

Line plot



## Theory

In a list of email addresses, you find a phone number. In the context of data quality, this is an issue of...

- *Accuracy*
- Completeness
- Credibility
- Understandability
- Precision