# 1   The evaluation scores for input data of various length for Unseen dataset

========================================
Comparing the evaluation scores between two model: (flat, modify baseline) according to different triple size in Unseen dataset.
*We mean by flat one model from the extended system which contains 3 model:flat,str1,str2*
*Where we evaluate just this model in term of different triple size, because the result of other models i.e. str1&str2 for the dataset seen and unseen not very good comparing by this model*
*We mean by "Modify baseline" the modified baseline of webnlg challeng*

## 1.1   Table of Evaluation in term of BLEU

| Triple size | Model | Bleu |
|---|---|---|
| | Modify Baseline | 16.62 |
| Triple 1 | Flat | **38.11** |
| | Modify Baseline | 9.15 |
| Triple 2 | Flat | **27.20** |
| | Modify Baseline | 10.56 |
| Triple 3 | Flat | **20.14** |
| | Modify Baseline | 10.96 |
| Triple 4 | Flat | **18.60** |
| | Modify Baseline | 7.72 |
| Triple 5 | Flat | **16.42** |

## 1.2   Table of Evaluation in term of METEOR

| Triple size | Model | METEOR |
|---|---|---|
| | Modify Baseline | 0.19 |
| Triple 1 | Flat | **0.41** |
| | Modify Baseline | 0.14 |
| Triple 2 | Flat | **0.34** |
| | Modify Baseline | 0.14 |
| Triple 3 | Flat | **0.28** |
| | Modify Baseline | 0.13 |
| Triple 4 | Flat | **0.26** |
| | Modify Baseline | 0.12 |
| Triple 5 | Flat | **0.23** |

## 1.3 Table of Evaluation in term of TER

| Triple size | Model | TER |
|---|---|---|
| | Modify Baseline | 0.76 |
| Triple 1 | Flat | **0.42** |
| | Modify Baseline | 0.92 |
| Triple 2 | Flat | **0.58** |
| | Modify Baseline | 0.89 |
| Triple 3 | Flat | **0.67** |
| | Modify Baseline | 0.88 |
| Triple 4 | Flat | **0.71** |
| | Modify Baseline | 0.91 |
| Triple 5 | Flat | **0.74** |