



Extracción de  
**vector de características de sesión**

```
def session_aggregation
```

**Preprocesado** para clustering:

**Muestreo** (solo para obtener  
datos de prueba)

```
$ shuf $ sort
```

**Agrupación** por día e IP origen  
Cálculo de **métricas**

```
./preprocess_clustering_dataset.py
```

**Estandarización**

**KMeans** con  $k=5$

