

TRABAJO PRÁCTICO 2 – SPARK

Nombre y Apellido: Mauricio Rodríguez Bertella

Padrón: 100624

Ayudante Asignado: Nati

Consulta 1: 1 -

<https://github.com/MauRodriguez/Orga-de-Datos-entregas/blob/master/TP%202/consulta1.ipynb>

Consulta 2: 8 -

<https://github.com/MauRodriguez/Orga-de-Datos-entregas/blob/master/TP%202/consulta8.ipynb>

CONSIDERACIONES: Tuve que tomar samples de pagelinks y de contents muy chicos porque si no no tenía forma de que se ejecutaran los comandos.

Consulta 3: 10 -

<https://github.com/MauRodriguez/Orga-de-Datos-entregas/blob/master/TP%202/consulta10.ipynb>

CONSIDERACIONES: Al hacer reduceByKey devuelve error TypeError: expected string or bytes-like object. Hablando con mi corrector probé distintas formas para solucionarlo. Al final pensamos quizás sería un caracter especial que molestaba, por eso filtré los caracteres especiales con re.sub, pero sigue sin andar. Finalicé la consulta sin poder compilar.

Consulta 4: 15 -

<https://github.com/MauRodriguez/Orga-de-Datos-entregas/blob/master/TP%202/consulta15.ipynb>

Consulta 5: 27 -

<https://github.com/MauRodriguez/Orga-de-Datos-entregas/blob/master/TP%202/consulta27.ipynb>

Consulta 6: 28 -

<https://github.com/MauRodriguez/Orga-de-Datos-entregas/blob/master/TP%202/consulta28.ipynb>