



Pandas: Split - Apply - Combine

75.06 / 95.58 Organización de Datos, 2020-1C

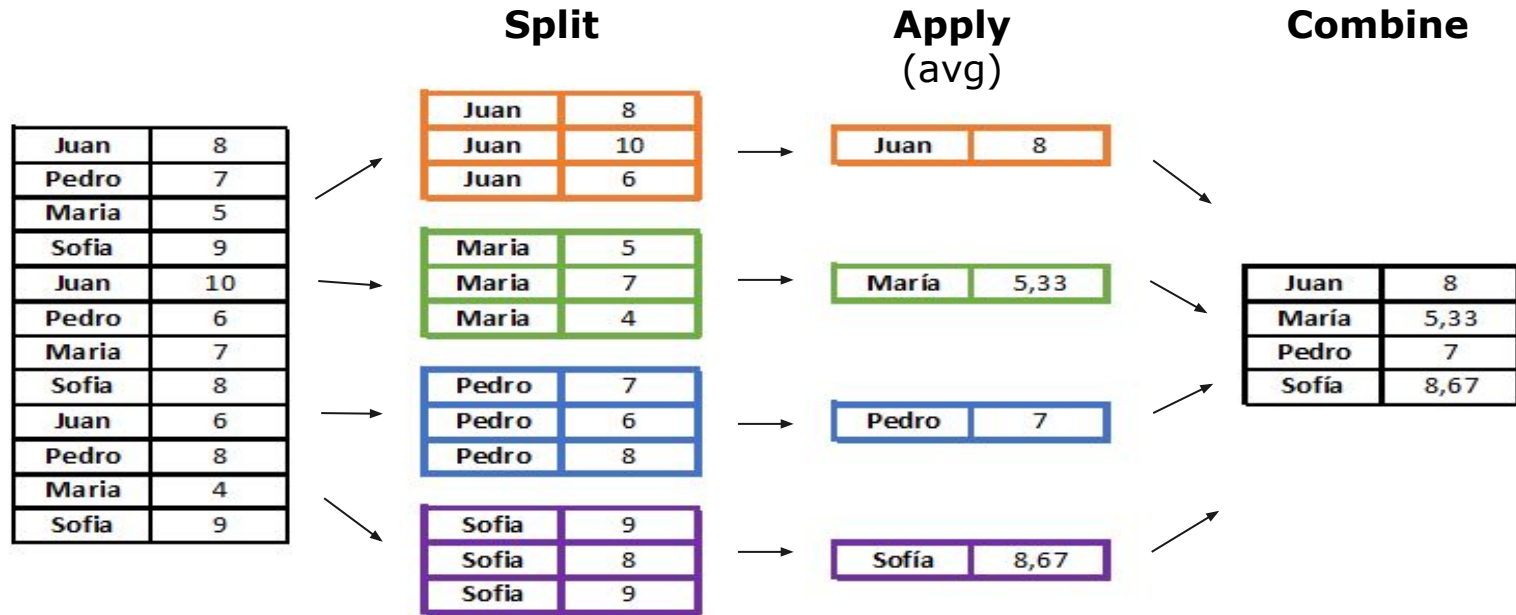


Motivación

Cuando manejamos datasets de datos, frecuentemente necesitamos analizar el comportamiento de ciertos grupos de datos

- Queremos calcular promedio de notas por curso de la facultad
- Queremos calcular las temperaturas máxima y mínima de las principales ciudades del país
- Queremos calcular los barrios con mayor precio promedio del m2 en las propiedades

Split - Apply - Combine





Split: GroupBy()

```
df.groupby('column_name')
```

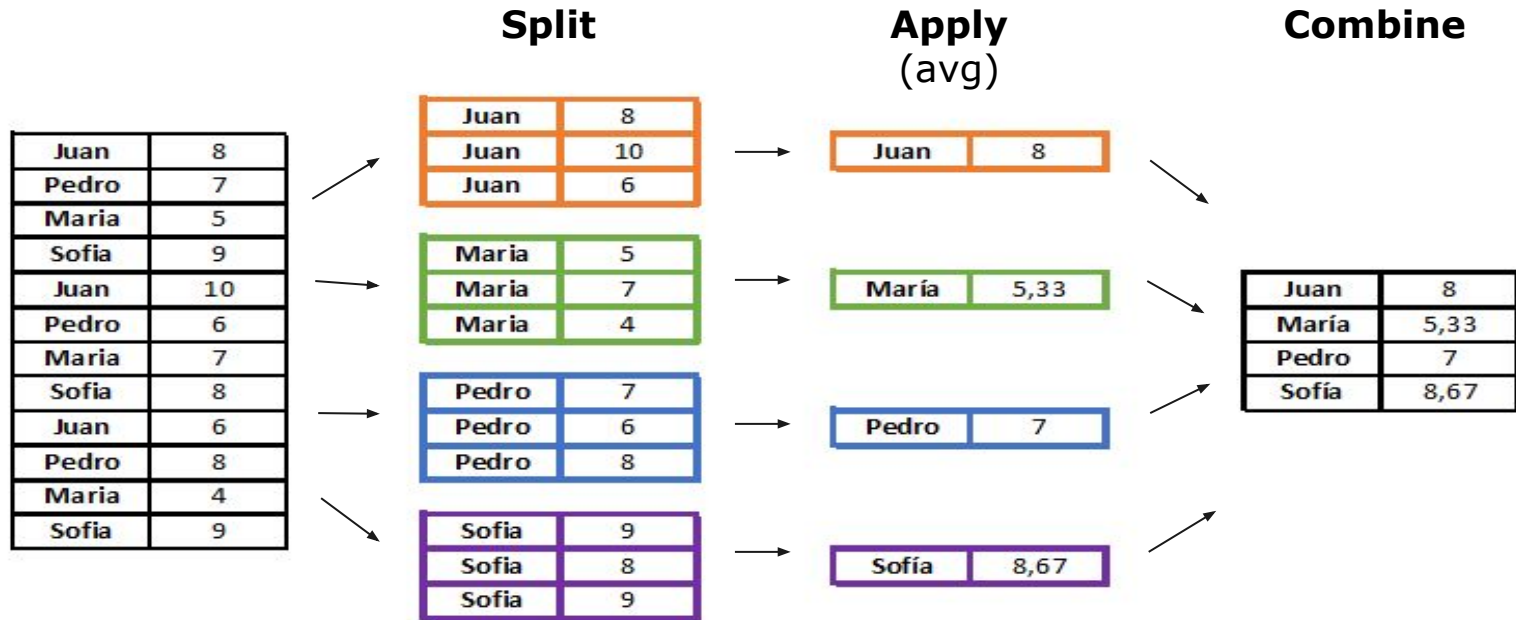
```
df.groupby(['lista', 'de', 'columnas'])
```



Devuelven un objeto del tipo Groupby

Sobre este objeto, podremos realizar agregaciones, filtrados, transformaciones y aplicar funciones (apply)

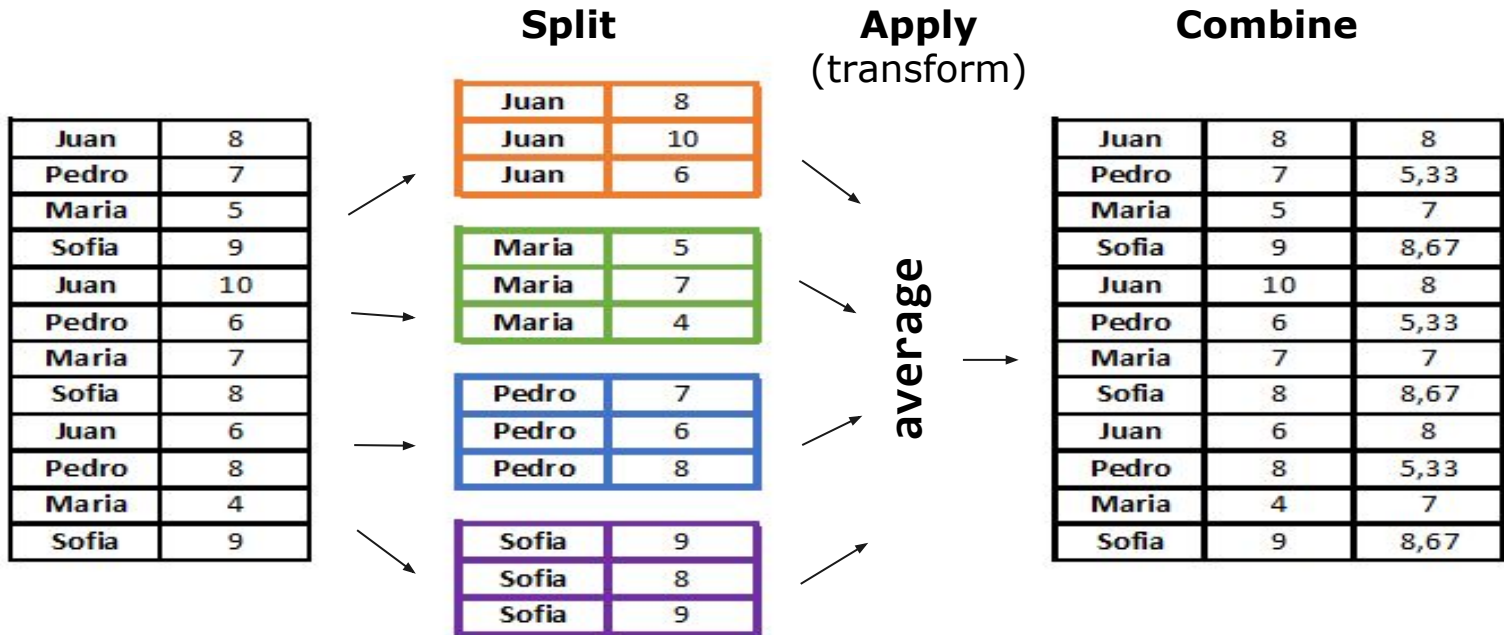
Agregaciones



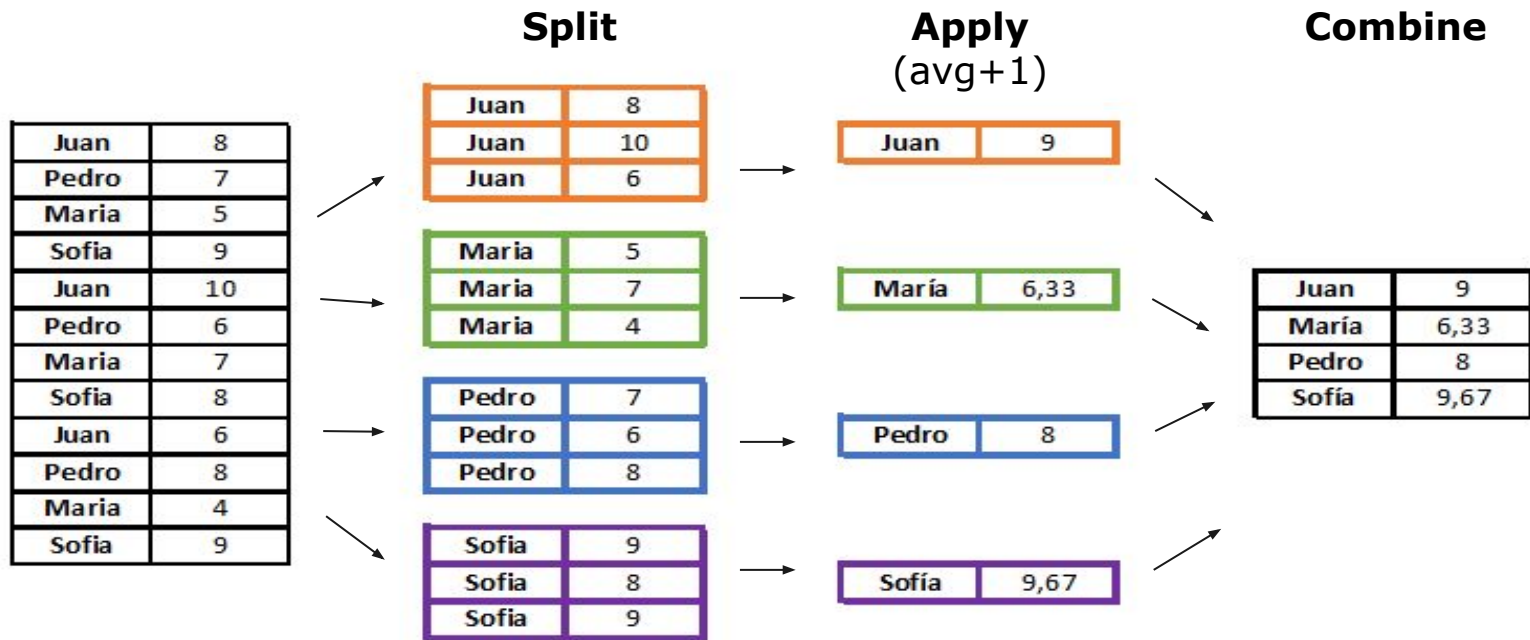
Filtrado



Transformaciones



Apply





Pandas

¡Pandas nos facilita resolver las operaciones antes mencionadas, en forma simple!