



Visión general: Modelo predictivo de métricas Stryd

David Víctor Gómez Ramírez

Técnico Superior en Desarrollo de Aplicaciones Multiplataforma especialidad en Big Data

[Escriba aquí una descripción breve del documento. Normalmente, una descripción breve es un resumen corto del contenido del documento. Escriba aquí una descripción breve del documento. Normalmente, una descripción breve es un resumen corto del contenido del documento.]

Contenido

INTRODUCCIÓN	1
STRYD	1
FUENTE DE LOS DATOS	4
PROCESOS DE ETL	6
TABLA STG_SEGMENTOS	7
MAPPINGS.....	7
WORFLOWS.....	8
VOLUMETRÍA Y EJECUCIÓN	9
FUENTES:.....	9

INTRODUCCIÓN

El objeto de este proyecto es la realización de un modelo de Machine Learning para la predicción de una/s métrica/s proporcionadas por un dispositivo que mide la potencia que generamos cuando realizamos la actividad física de correr. Como podemos observar esta contextualización es bastante vaga e indefinida, pero como todo proyecto debemos empezar por algún punto. Vayamos pues a ello.

STRYD

¿Qué es STRYD?

Stryd es un medidor de la potencia que generamos cuando practicamos “running” y, aunque en concepto es muy similar a los medidores de potencia del “ciclismo”, tienen diferencias sustanciales ya que Stryd no mide la potencia de forma directa sino que utiliza un algoritmo para su cálculo.

Para este proyecto personal, Stryd no es más que una herramienta más, por lo que no ahondaremos en su funcionamiento pero sí en las métricas y datos que nos aporta y que serán la base de nuestro modelo predictivo, es decir, será la fuente principal de la que obtendremos los datos de nuestro modelo.

Por simplificar y ubicarnos mejor en el contexto diremos que Stryd es un dispositivo que mide diferentes valores cuando realizamos entrenamientos de “running” y cuyo eje central es la potencia medida en watos, pero tenemos muchas más y que serán muy importantes para nuestro modelo:

- **Tiempo (DURACION):** variable fundamental de cualquier entrenamiento y que mide la duración del mismo.
- **Distancia (DISTANCIA):** mide la longitud de los entrenamientos.
- **Ritmo (RITMO):** es la métrica por excelencia que relaciona “Tiempo” y “Distancia”.

Visión general: Modelo predictivo de métricas Stryd



- **Altimetría y Desnivel (ALTITUD – DESNIVEL):** mediante mide la altitud y sus variaciones.



- **Cadencia (CADENCIA):** son las zancadas o pasos que damos mientras corremos y se mide en zancadas o pasos por minuto.
- **TSC (Tiempo de contacto con el suelo):** podríamos definirlo como la duración medida en ms en el que nuestro pie está en contacto con el suelo.



Visión general: Modelo predictivo de métricas Stryd

- **Oscilación vertical (OSC_VERTICAL):** es la distancia vertical que recorre el cuerpo desde el punto en que el centro de gravedad está más bajo hasta el punto más alto.
- **Leg Spring Stiffness (LSS):** se trataría del acrónimo inglés LSS (Leg Spring Stiffness) que adaptado al castellano vendría a ser cómo la rigidez del resorte.



- **Pendiente (PENDIENTE):** mide el porcentaje de pendiente media durante la distancia recorrida.
- **RSS (RSS):** mide el estrés que ha generado el entrenamiento o segmento.
- **Potencia (POTENCIA):** mide la potencia media generada durante un entrenamiento o segmento.
- **Frecuencia Cardíaca (FREC_CARDIACA):** mide la frecuencia cardíaca por entrenamiento o segmento medida en pulsaciones por minuto (ppm)
- **Aire (AIRE):** mide la resistencia de aire que encontramos en un entrenamiento o segmento.
- **Form Power (FP):** mide la potencia vertical y lateral que generamos cuando nos desplazamos en un entrenamiento o segmento.



- **Segmento:** es un concepto que deberemos tener muy presente y que no es más que pequeñas partes del entrenamiento de los que podemos extraer la misma información.

Visión general: Modelo predictivo de métricas Stryd

Segmento	Tiempo en mov.	Distancia	Potencia	Ritmo	Cadencia	FC	Potencia vertical	BPV	TCS	PP	OV	Resistencia del aire
1	8:52	757 m	180 W	9:04 /km	158 spm	123 bpm	87 W	0.42	343 ms	12.4 kN/m	5.83 cm	0%
2	8:44	748 m	183 W	8:58 /km	160 spm	122 bpm	88 W	0.42	342 ms	12.3 kN/m	5.73 cm	0%
3	8:44	1.00 km	185 W	8:44 /km	160 spm	138 bpm	88 W	0.40	343 ms	11.9 kN/m	5.82 cm	0%
4	8:44	1.00 km	187 W	8:43 /km	160 spm	144 bpm	88 W	0.40	342 ms	11.9 kN/m	5.74 cm	0%
5	8:29	999 m	170 W	8:30 /km	160 spm	149 bpm	89 W	0.40	338 ms	11.9 kN/m	5.87 cm	0%
6	8:33	750 m	185 W	8:44 /km	161 spm	150 bpm	89 W	0.40	341 ms	11.9 kN/m	5.89 cm	0%
7	8:34	743 m	184 W	8:51 /km	161 spm	152 bpm	87 W	0.40	341 ms	12.1 kN/m	5.83 cm	0%

FUENTE DE LOS DATOS

La herramienta Stryd dispone de su propia aplicación móvil y de su equivalente de escritorio en formato web, para poder extraer los datos de los entrenamientos deberemos descargarnos los archivos desde la aplicación móvil en formato csv. Por cada entrenamiento deberemos descargarnos 3 archivos aunque en principio sólo utilizaremos 2. Un archivo en los que las distancias de los segmentos se han definido manualmente y un segundo donde las distancia de los segmentos está predefinida a 1000m.

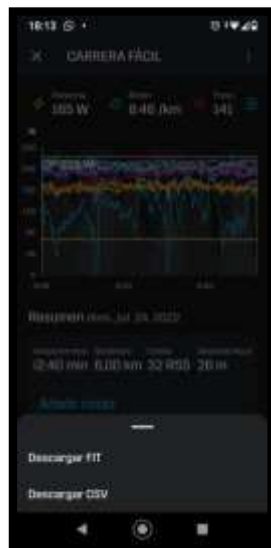
Manualmente aplicación web

Tabla de segmentos: Manual		Distancia	Potencia	Ritmo	Cadencia	FC	Potencia vertical	BPV	TCS	PP	OV	Resistencia del aire
Segmento	Tiempo en mov.	Distancia	Potencia	Ritmo	Cadencia	FC	Potencia vertical	BPV	TCS	PP	OV	Resistencia del aire
1	8:52	757 m	180 W	9:04 /km	158 spm	123 bpm	87 W	0.42	343 ms	12.4 kN/m	5.83 cm	0%
2	8:44	748 m	183 W	8:58 /km	160 spm	122 bpm	88 W	0.42	342 ms	12.3 kN/m	5.73 cm	0%
3	8:44	1.00 km	185 W	8:44 /km	160 spm	138 bpm	88 W	0.40	343 ms	11.9 kN/m	5.82 cm	0%
4	8:44	1.00 km	187 W	8:43 /km	160 spm	144 bpm	88 W	0.40	342 ms	11.9 kN/m	5.74 cm	0%
5	8:29	999 m	170 W	8:30 /km	160 spm	149 bpm	89 W	0.40	338 ms	11.9 kN/m	5.87 cm	0%
6	8:33	750 m	185 W	8:44 /km	161 spm	150 bpm	89 W	0.40	341 ms	11.9 kN/m	5.89 cm	0%
7	8:34	743 m	184 W	8:51 /km	161 spm	152 bpm	87 W	0.40	341 ms	12.1 kN/m	5.83 cm	0%

Predefinidos aplicación web

Tabla de segmentos: Manual		Distancia	Potencia	Ritmo	Cadencia	FC	Potencia vertical	BPV	TCS	PP	OV	Resistencia del aire
Segmento	Tiempo en mov.	Distancia	Potencia	Ritmo	Cadencia	FC	Potencia vertical	BPV	TCS	PP	OV	Resistencia del aire
1	8:53	1.00 km	181 W	8:59 /km	158 spm	123 bpm	87 W	0.42	342 ms	12.3 kN/m	5.89 cm	0%
2	8:50	1.00 km	184 W	8:50 /km	160 spm	134 bpm	88 W	0.40	342 ms	12.2 kN/m	5.75 cm	0%
3	8:43	1.00 km	188 W	8:43 /km	160 spm	140 bpm	88 W	0.40	343 ms	11.8 kN/m	5.80 cm	0%
4	8:41	1.00 km	188 W	8:41 /km	160 spm	146 bpm	88 W	0.40	340 ms	12.0 kN/m	5.80 cm	0%
5	8:34	1.00 km	188 W	8:34 /km	161 spm	150 bpm	88 W	0.40	338 ms	11.9 kN/m	5.75 cm	0%
6	8:53	999 m	184 W	8:54 /km	161 spm	152 bpm	87 W	0.40	342 ms	12.1 kN/m	5.85 cm	0%

Visión general: Modelo predictivo de métricas Stryd



Segm ento	Duración	Distancia	Potencia	Ritmo	Frecuencia cardíaca	Cadencia
1	8:59 min	1,00 km	161 W	8:59 /km	123 bpm	158 ppm
2	8:51 min	1,00 km	164 W	8:51 /km	134 bpm	160 ppm
3	8:43 min	1,00 km	166 W	8:43 /km	140 bpm	160 ppm
4	8:40 min	1,00 km	168 W	8:40 /km	146 bpm	160 ppm
5	8:34 min	1,00 km	168 W	8:34 /km	150 bpm	161 ppm
6	8:53 min	1,00 km	164 W	8:54 /km	152 bpm	161 ppm

Segm ento	Duración	Distancia	Potencia	Ritmo	Frecuencia cardíaca	Cadencia
1	6:52 min	0,76 km	160 W	9:04 /km	121 bpm	158 ppm
2	6:44 min	0,75 km	163 W	8:59 /km	132 bpm	160 ppm
3	8:44 min	1,00 km	165 W	8:44 /km	136 bpm	160 ppm
4	8:44 min	1,00 km	167 W	8:43 /km	144 bpm	160 ppm
5	8:29 min	1,00 km	170 W	8:29 /km	149 bpm	160 ppm
6	6:33 min	0,75 km	165 W	8:44 /km	150 bpm	161 ppm
7	6:34 min	0,74 km	164 W	8:50 /km	152 bpm	161 ppm

Mediante la herramienta Powercenter de Stryd podremos obtener los datos de los segmentos que queramos y así poder obtener más información.

Segmento	Duración	Distancia	Potencia	Ritmo	Frecuencia cardíaca	Cadencia	Potencia vertical	Ratio de potencia vertical	Ratio de potencia horizontal	Ratio de potencia total	Ratio de potencia de contacto	Ratio de potencia de apoyo	Ratio de potencia de propulsión	Ratio de potencia de freno	Ratio de potencia de desaceleración	Ratio de potencia de aceleración	Ratio de potencia de desaceleración	Ratio de potencia de aceleración	Ratio de potencia de desaceleración	Ratio de potencia de aceleración
1	8:57 min	1,00 km	160 W	8:57 /km	123 bpm	159 ppm	67 W	0,42	135 ms	12,4 kN/m	5,66 cm	0 %	0,68 m	0,1 %	25 m	5 R55	4 m	5 m	5 m	5 m
2	8:40 min	1,00 km	165 W	8:40 /km	120 bpm	162 ppm	68 W	0,41	136 ms	12,3 kN/m	5,64 cm	0 %	0,71 m	0,1 %	24 m	5 R55	4 m	5 m	5 m	5 m
3	8:16 min	1,00 km	174 W	8:16 /km	127 bpm	164 ppm	68 W	0,39	125 ms	12,1 kN/m	5,67 cm	0 %	0,74 m	0,1 %	24 m	6 R55	5 m	5 m	5 m	5 m
4	8:03 min	1,00 km	179 W	8:03 /km	132 bpm	164 ppm	69 W	0,39	125 ms	11,8 kN/m	5,77 cm	0 %	0,76 m	0,1 %	25 m	7 R55	5 m	5 m	5 m	5 m
5	8:00 min	1,00 km	179 W	8:00 /km	136 bpm	164 ppm	69 W	0,38	126 ms	11,9 kN/m	5,71 cm	0 %	0,75 m	0,2 %	25 m	7 R55	5 m	5 m	5 m	5 m
6	8:03 min	1,00 km	177 W	8:03 /km	140 bpm	164 ppm	69 W	0,39	126 ms	11,9 kN/m	5,69 cm	0 %	0,76 m	0,2 %	25 m	6 R55	4 m	5 m	5 m	5 m
7	8:27 min	1,00 km	172 W	8:27 /km	143 bpm	163 ppm	68 W	0,4	132 ms	12,1 kN/m	5,63 cm	0 %	0,72 m	0,1 %	25 m	6 R55	5 m	5 m	5 m	5 m
8	8:43 min	1,00 km	166 W	8:44 /km	142 bpm	163 ppm	67 W	0,41	133 ms	12,3 kN/m	5,56 cm	0 %	0,7 m	0 %	25 m	5 R55	4 m	5 m	5 m	5 m

Visión general: Modelo predictivo de métricas Stryd

[illegible][illegible]

Tras unificar **“SEGMENTOS1K”** y **“SEGMENTOS_MANUAL”** realizaremos algunas transformaciones para adaptar los datos a nuestras necesidades en un archivo de texto plano **“SEGMENTOS.txt”**

[illegible]

PROCESOS DE ETL

Nos podríamos extender y podríamos debatir sobre qué herramienta es mejor o peor, nos gusta más o menos o es más apropiada o menos apropiada para las tareas de ETL. Particularmente y por su sencillez utilizaremos Informatica PowerCenter y como base de datos SQL SERVER sin despreciar otras posibles herramientas y soluciones.

Visión general: Modelo predictivo de métricas Stryd

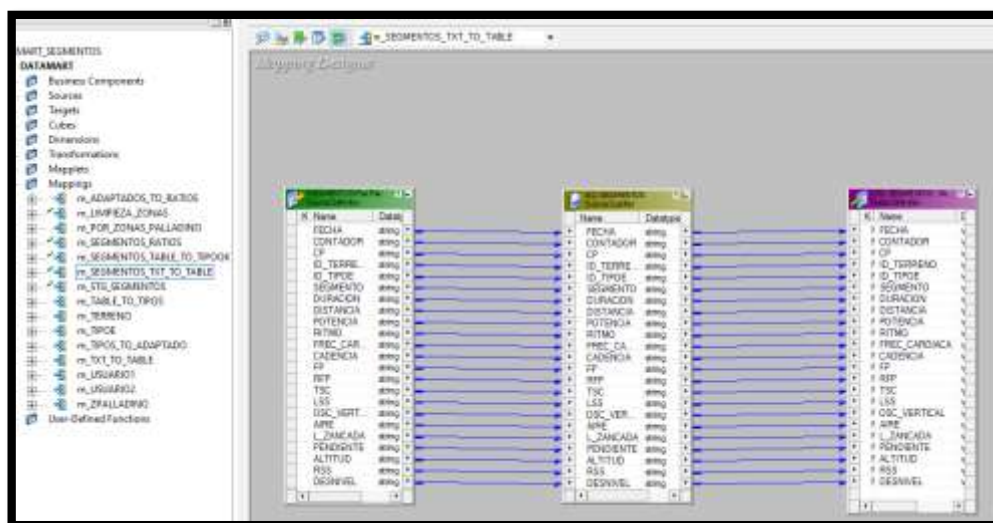
Deberemos crear diferentes “mappings” y otros tantos “workflows” con el objetivo de llevar nuestro archivo de texto plano “SEGMENTOS.txt” hasta la tabla “STG_SEGMENTOS”, realizando conversiones de tipos, transformaciones, operaciones, filtros ,etc....

TABLA STG_SEGMENTOS

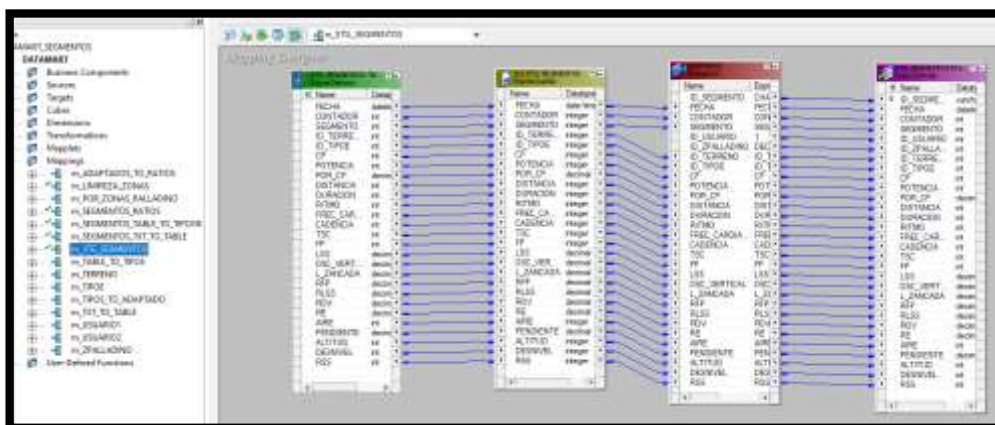
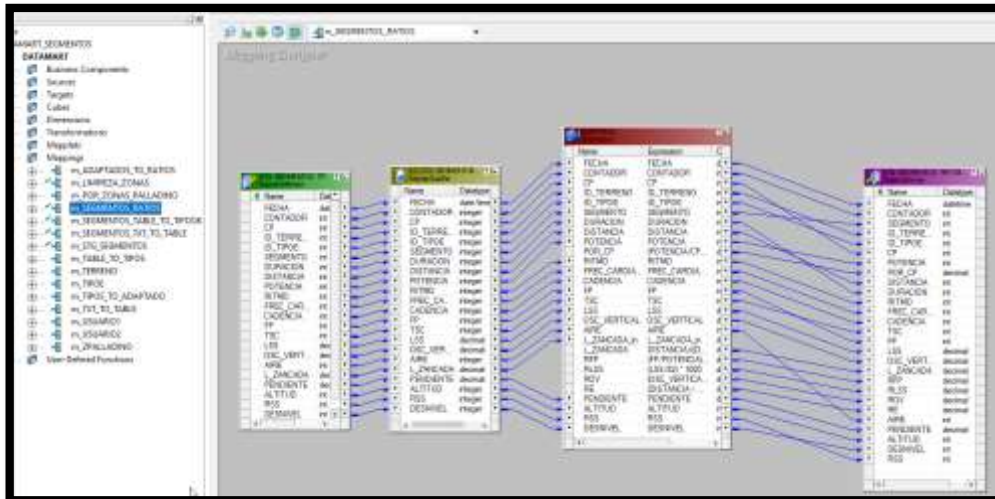
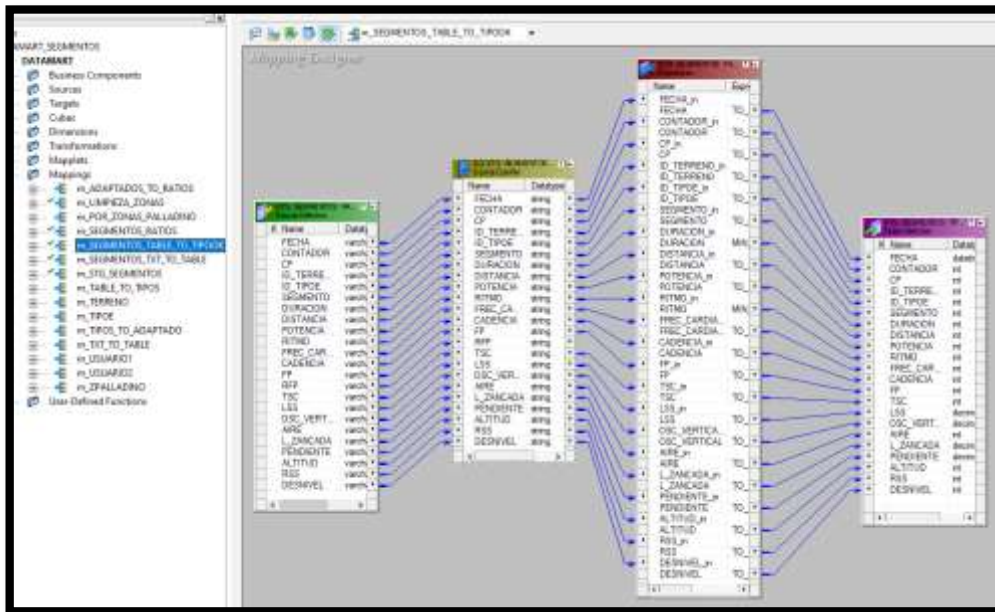
```
CREATE TABLE [dbo].[STG_SEGMENTOS]
(
    [ID_SEGMENTO] [varchar](20) NOT NULL,
    [FECHA] [date] NULL,
    [CONTADOR] [int] NULL,
    [SEGMENTO] [int] NULL,
    [ID_USUARIO] [int] NULL,
    [ID_ZONAS_PALLADINO] [int] NULL,
    [ID_TERRENO] [int] NULL,
    [ID_TIPO] [int] NULL,
    [CP] [int] NULL,
    [POTENCIA] [int] NULL,
    [POR_CP] [decimal](18, 3) NULL,
    [DISTANCIA] [int] NULL,
    [DURACION] [int] NULL,
    [RITMO] [int] NULL,
    [FREC_CARDIACA] [int] NULL,
    [CAIDENCIA] [int] NULL,
    [TSC] [int] NULL,
    [LSS] [int] NULL,
    [OSC_VERTICAL] [decimal](18, 8) NULL,
    [L_ZANCADEA] [decimal](18, 3) NULL,
    [APP] [decimal](18, 3) NULL,
    [RUS] [decimal](18, 3) NULL,
    [RWD] [decimal](18, 3) NULL,
    [RE] [decimal](18, 3) NULL,
    [ADRC] [int] NULL,
    [PONDIENTE] [decimal](18, 8) NULL,
    [ALTITUD] [int] NULL,
    [OPINION] [int] NULL,
    [RSS] [int] NULL,
    PRIMARY KEY CLUSTERED
(
    [ID_SEGMENTO] ASC
)
WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) IN [PRIMARY]
) ON [PRIMARY]
```

ID_SEGMENTO	FECHA	CONTADOR	SEGMENTO	ID_USUARIO	ID_ZALLADINO	ID_TERRENO	ID_TIPO	CP	POTENCIA	POR_CP	DISTANCIA	DURACION	RITMO	FREC_CARDIACA	CAIDENCIA	TSC
20224110372191164	2022-04-11	10	3	1	5	7	2	191	164	85.963	500	259	519	154	165	24
20224111172191162	2022-04-11	1	1	1	5	7	2	191	162	84.816	1000	547	547	130	189	24
20224112272191161	2022-04-11	2	2	1	5	7	2	191	161	84.293	1000	537	537	137	181	24
20224113372191162	2022-04-11	3	3	1	5	7	2	191	162	84.816	1000	540	540	142	162	24
20224114472191162	2022-04-11	4	4	1	5	7	2	191	162	84.816	1000	539	539	146	183	24
20224115572191162	2022-04-11	5	5	1	5	7	2	191	162	84.816	1000	542	542	150	183	24
20224116672191162	2022-04-11	6	6	1	5	7	2	191	162	84.816	1000	537	537	152	195	23
20224117772191164	2022-04-11	7	7	1	5	7	2	191	164	85.963	600	317	524	155	165	24
20224118872191163	2022-04-11	8	1	1	5	7	2	191	163	85.940	520	300	561	128	150	24
20224119972191162	2022-04-11	9	2	1	5	7	2	191	162	84.816	5570	3008	538	145	183	24
2022411949472191162	2022-04-11	94	94	1	5	7	2	191	162	84.816	2000	1082	545	134	180	24
2022411959572191162	2022-04-11	95	95	1	5	7	2	191	162	84.816	2000	1080	540	144	162	24
2022411969672191162	2022-04-11	96	96	1	5	7	2	191	162	84.816	2000	1079	538	151	164	24
202241197972191162	2022-04-11	97	97	1	5	7	2	191	162	84.816	3000	1622	541	136	181	24
2022411989872191162	2022-04-11	98	98	1	5	7	2	191	162	84.816	3000	1619	539	149	184	24
2022411999972191162	2022-04-11	99	99	1	5	7	2	191	162	84.816	6610	3859	539	144	162	24
20224120112191166	2022-04-12	1	1	1	5	1	2	191	166	86.910	1000	520	520	126	189	23
20224120212191177	2022-04-12	2	2	1	6	1	2	191	177	92.670	1000	496	496	137	184	23
20224123012191198	2022-04-12	3	3	1	8	1	2	191	196	103.664	1000	432	432	151	171	28

MAPPINGS



Visión general: Modelo predictivo de métricas Stryd



WORKFLOWS

Los flujos de trabajo que estamos desarrollando para nuestro modelo predictivo sería los 4 de abajo, que son los que realizan las transformaciones necesarias hasta llevarlos a la tabla "STG_SEGMENTOS"

Visión general: Modelo predictivo de métricas Stryd



VOLUMETRÍA Y EJECUCIÓN

The screenshot shows the Informatica PowerCenter Workflow Monitor. The 'Workflow Run' table lists the execution of the 'wf_ACTUALIZACIONES' workflow. The table has columns for 'Start Time', 'Completion Time', and 'Status'. A red box highlights the 'i_m_SEGMENTS' task, which is shown as 'Succeeded'.

Task Name	Start Time	Completion Time	Status
wf_ACTUALIZACIONES	02/08/2022 19:02:31	02/08/2022 19:06:06	Succeeded
i_m_USUARIO1	02/08/2022 19:02:35	02/08/2022 19:02:56	Succeeded
i_m_USUARIO2	02/08/2022 19:02:59	02/08/2022 19:03:16	Succeeded
i_m_ZPALLADINO	02/08/2022 19:03:16	02/08/2022 19:03:33	Succeeded
i_m_TIPOE	02/08/2022 19:03:34	02/08/2022 19:03:50	Succeeded
i_m_TERRENO	02/08/2022 19:03:51	02/08/2022 19:04:08	Succeeded
i_m_SEGMENTS_TXT_TO_TABLE	02/08/2022 19:04:08	02/08/2022 19:04:25	Succeeded
i_m_SEGMENTS_TABLE_TO_T...	02/08/2022 19:04:26	02/08/2022 19:04:42	Succeeded
i_m_SEGMENTS_RATIOS	02/08/2022 19:04:43	02/08/2022 19:04:59	Succeeded
i_m_STG_SEGMENTS	02/08/2022 19:05:00	02/08/2022 19:05:16	Succeeded

Transformation Name	Node	Applied Rows	Rejected Rows	Throughput (Rows/Sec)	Throughput
sq_STG_SEGMENT...	node01_DESKT...	1229	0	1229	191724
STG_SEGMENTS	node01_DESKT...	1229	0	1229	227365

FUENTES:

- Eduard Barceló: <https://www.eduardbarcelo.com/metricas-de-stryd/>
- Correr una Maratón: <https://www.correrunamaraton.com/stryd-medidor-potencia/>

