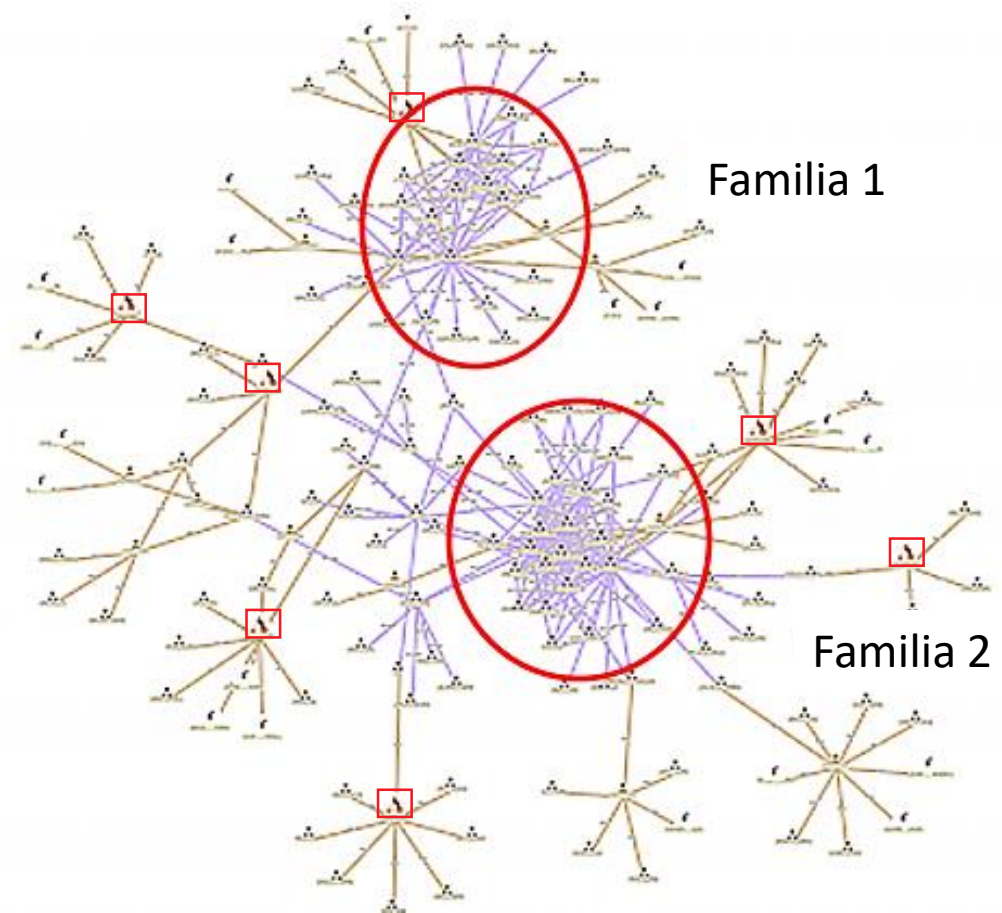
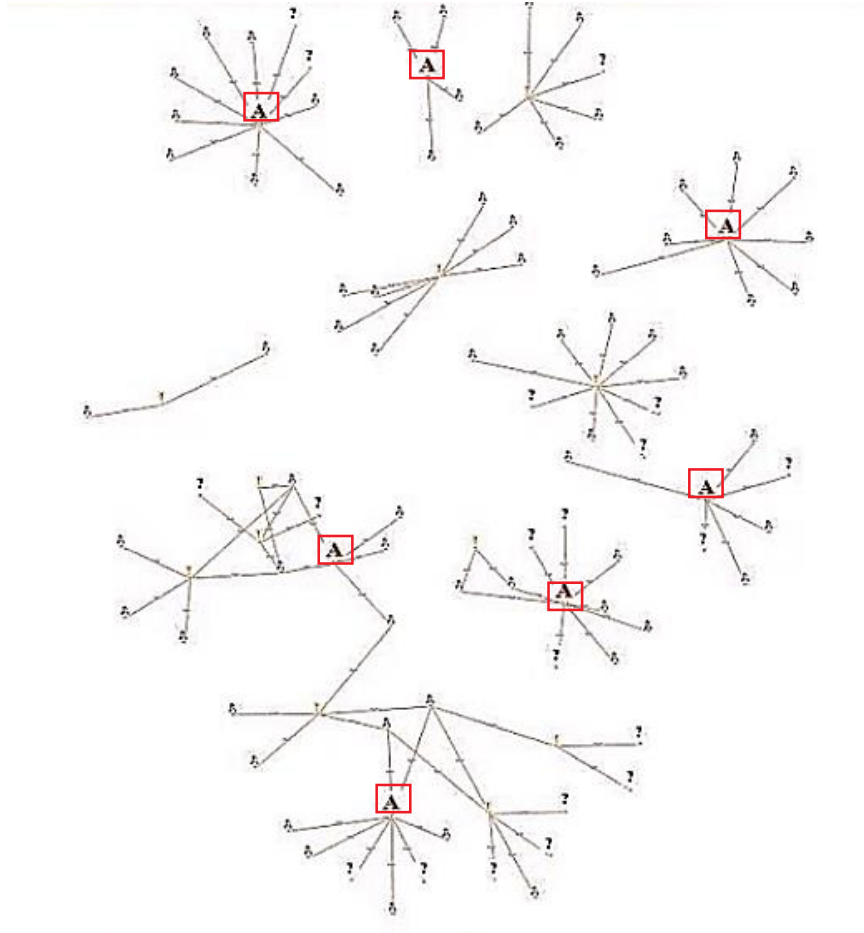






Dominando Big Data con HPCC Systems

Descripción general y aplicaciones de la plataforma

Caso de uso: Fraude de seguros



-  Accidentes con sospecha de fraude
-  Personas asociadas a accidentes

¡Bienvenido! - Agenda del taller

- ✓ HPCC Systems: descripción general
 - ✓ ¿Qué es? ¿De donde vino? ¿Para que sirve?
- ✓ Tutorial: ETL con ECL
 - ✓ Hands-on
- ✓ Bundles y aplicaciones
 - ✓ Visualización
 - ✓ PLN
 - ✓ Machine Learning

Recursos del taller

- Computadora personal (ECL IDE v.7.6.64 / VSCode)
- Clúster HPCC Systems: <http://3.139.124.33:8010/>
- Diapositivas y códigos (github.com/hpccsystems-solutions-lab/hpcc-systems-BR)
- Certificado de participación

¿Quienes somos?

- ✓ hugo.watanuki@lexisnexisrisk.com
- ✓ robert.foreman@lexisnexisrisk.com
- ✓ richard.taylor@lexisnexisrisk.com
- ✓ <https://hpccsystems.com/bb/>

Meet the Trainers



Richard Taylor

Richard is the original author of the ECL documentation, developer and designer of the HPCC Systems Training Courses, and is the Chief Instructor for all classroom and remote based training.

Bob Foreman

Bob is the developer and designer of the HPCC Systems Online Training Courses, and is the Senior Instructor for all classroom and online based training.

Hugo Watanuki

Hugo supports the development and delivery of training programs for the HPCC Systems platform in the Brazil region.

HPCC Systems: descripción general

¿Qué es?

HPCC Systems (*High Performance Computing Cluster*) es una plataforma para resolver desafíos de Big Data:

- **Supercomputación:** procesamiento paralelo y datos distribuidos
- **Open source:** libre y de código abierto
- **Completa:** gestión de flujo de datos completa y simplificada

¿De donde vino?

2001



Se lanza la
primera versión

2011



Código abierto (licencia
Apache y código en
GitHub)

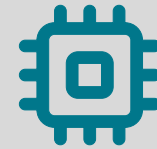
2012 – 16



Mejoras continuas con
foco en la calidad

Capacitación y soporte
mejorados

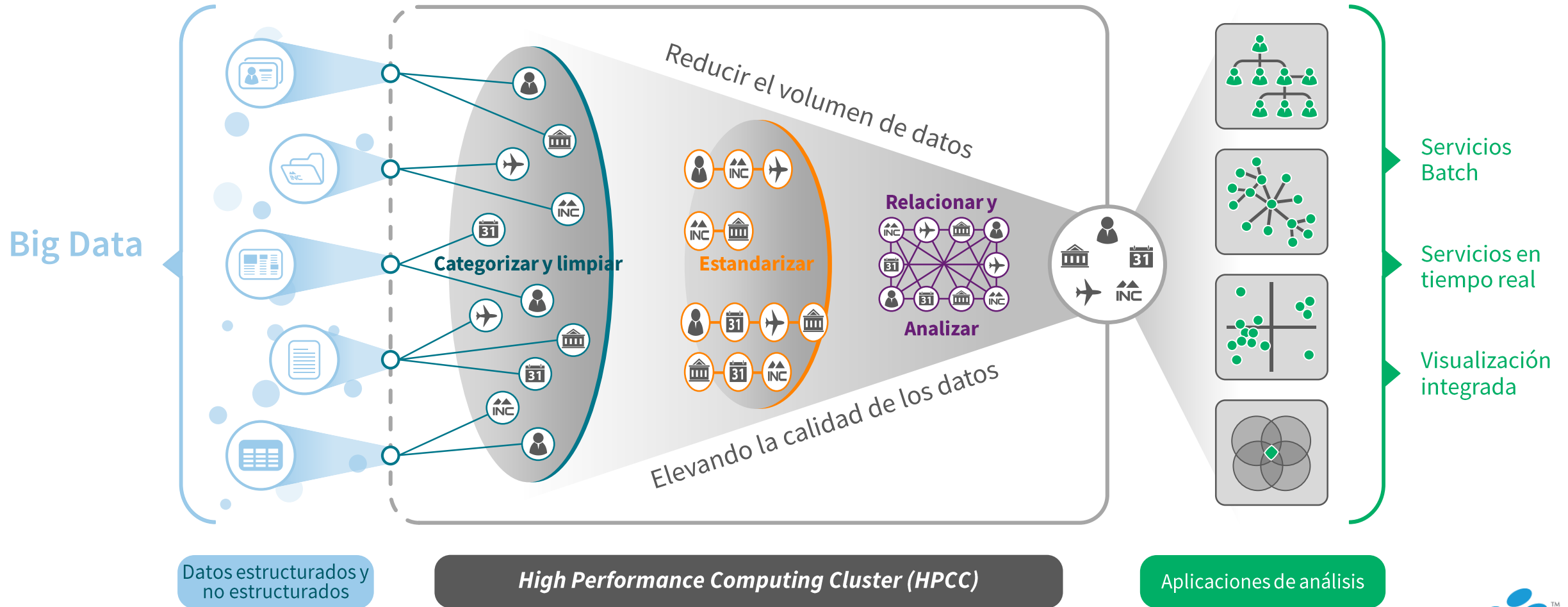
2017-actualidad



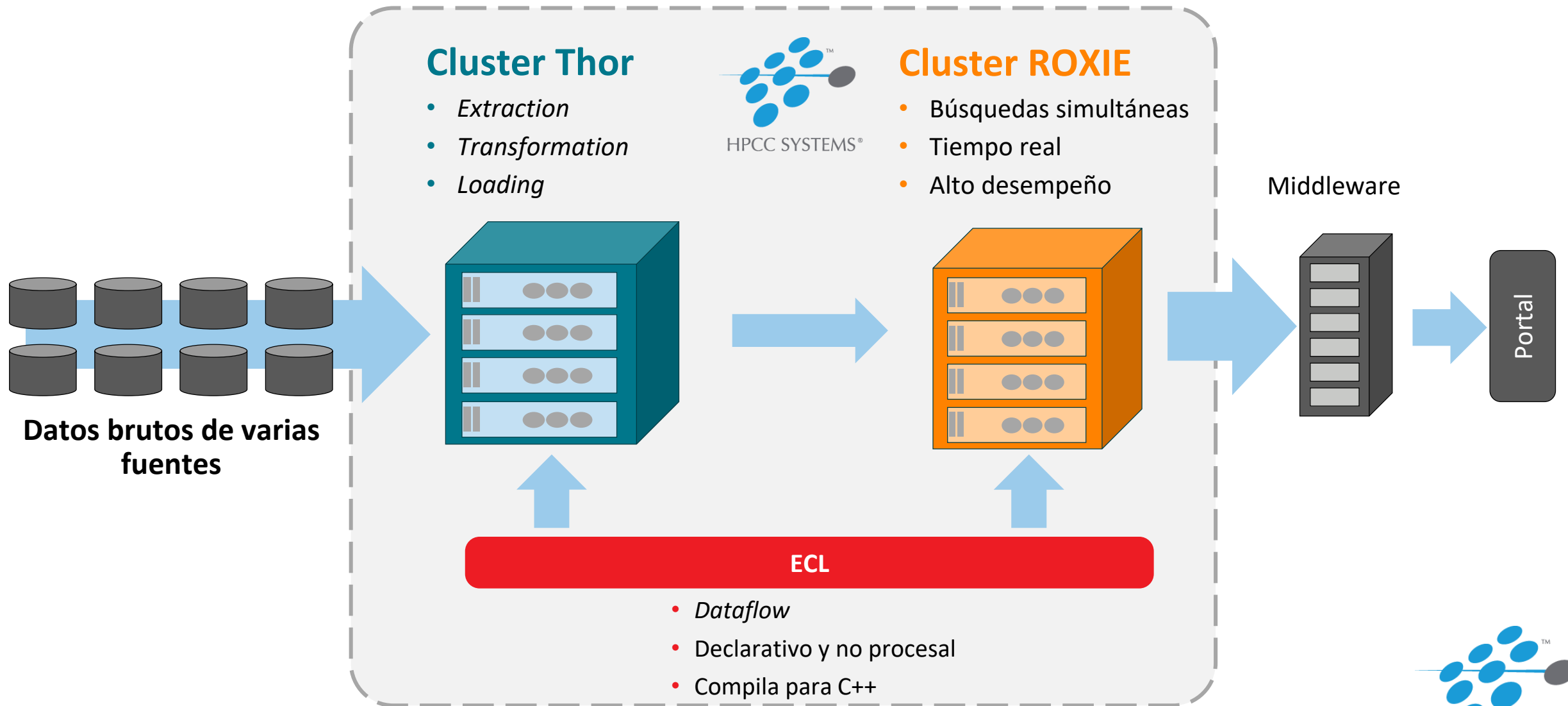
Mejoras arquitectónicas
(Cloud)

Desarrollos en Machine
Learning

¿Para que sirve?



El *power trio* de la plataforma: Thor, ROXIE y ECL



Thor y ROXIE: objetivos específicos

Thor:
*“Identificar y
catalogar todos los
seres vivos de los
océanos.”*

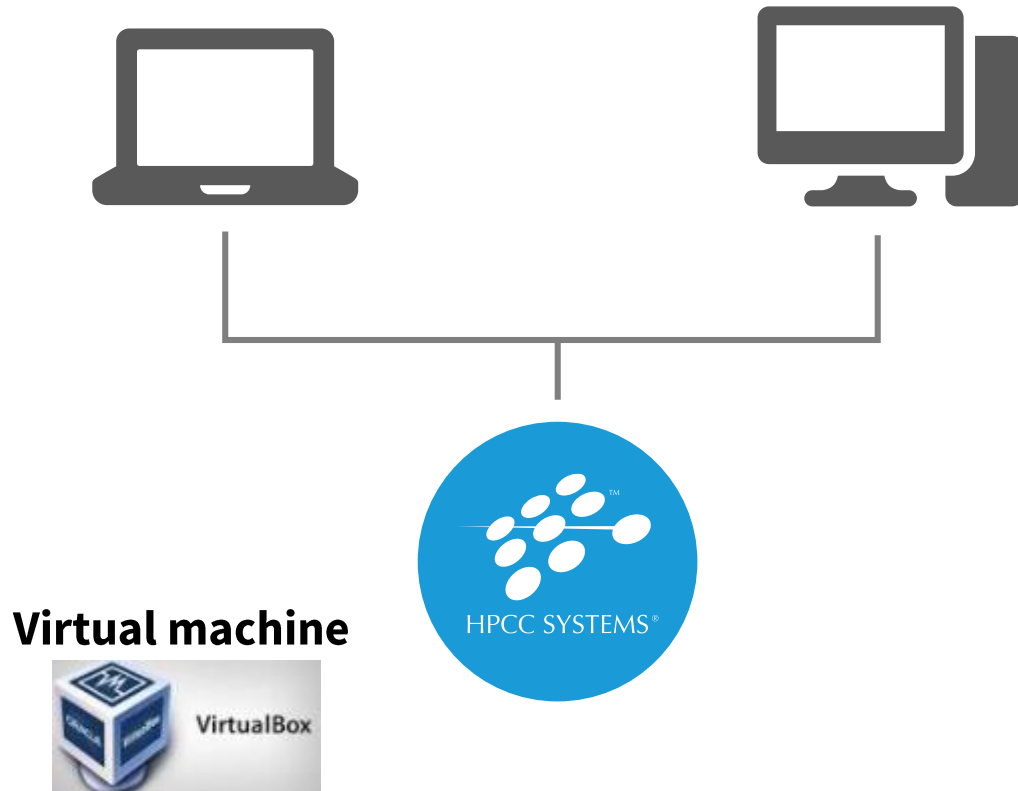


ROXIE:
*“Poner a disposición
toda la información
sobre una especie”*



La plataforma puede funcionar en ...

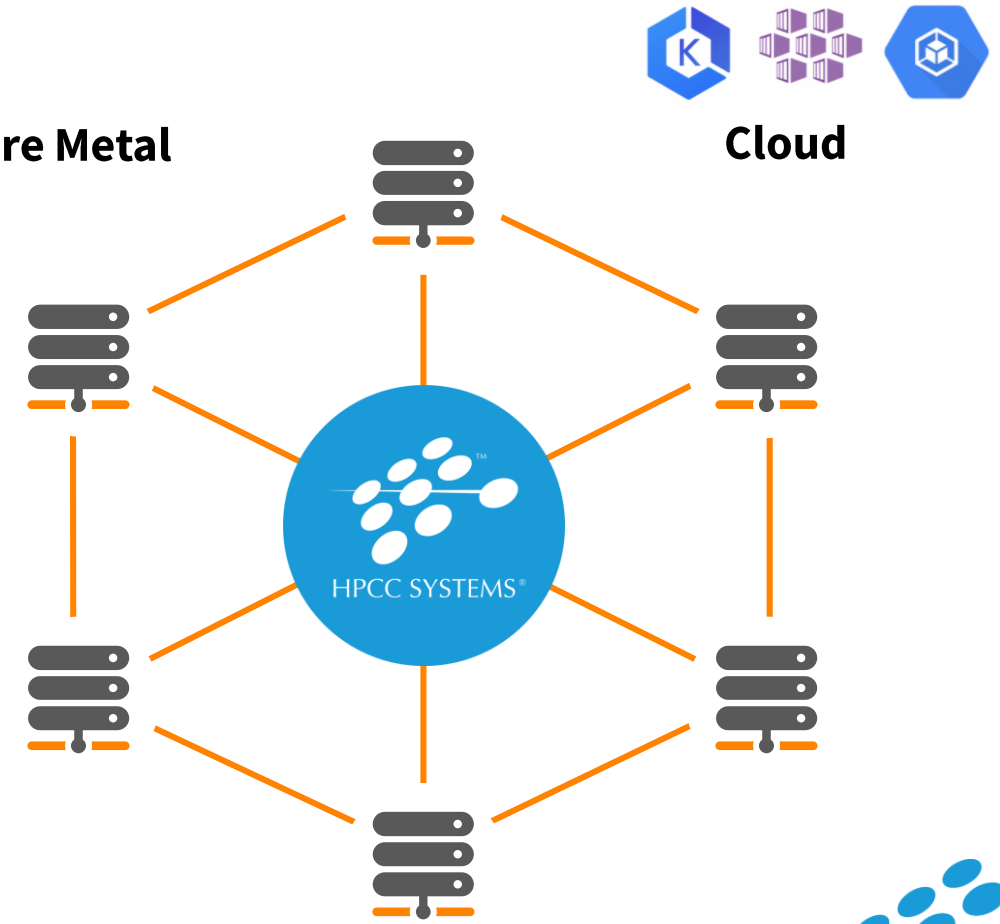
... una sola computadora.



... un clúster.

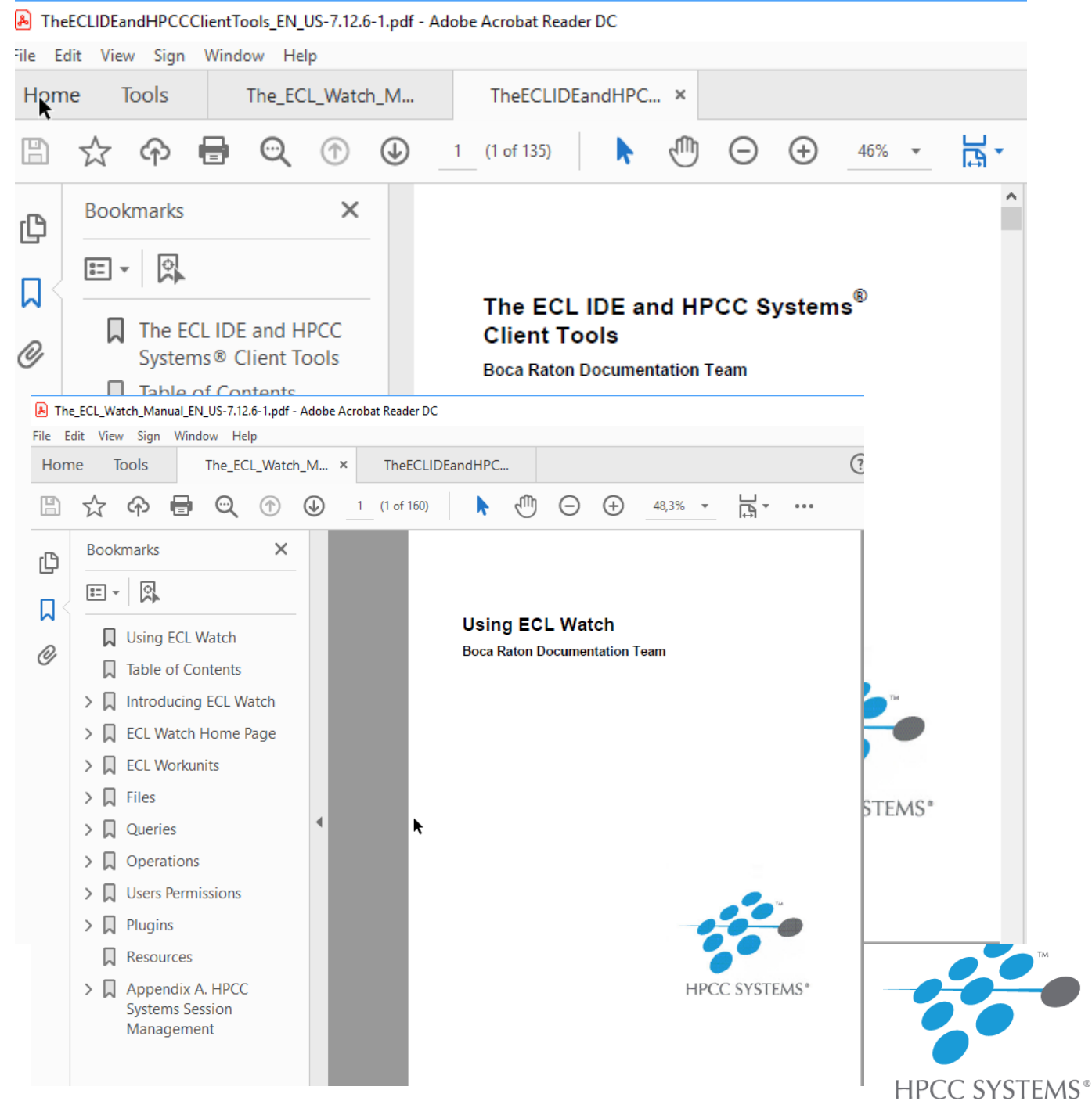
Bare Metal

Cloud



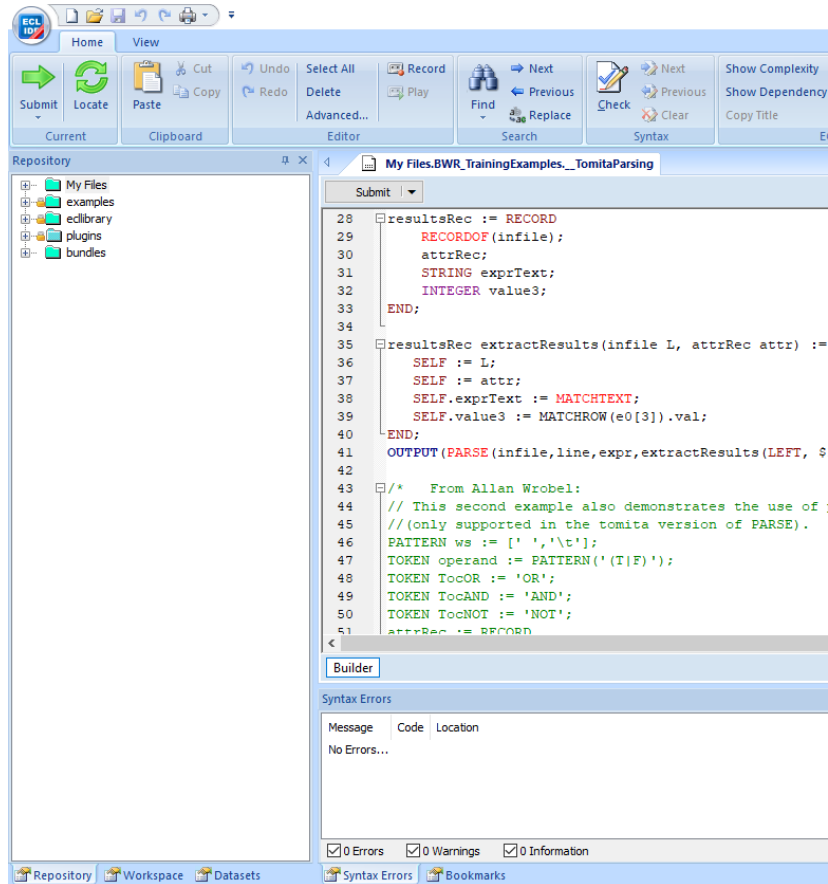
Interfaces de uso

- ✓ ECL IDE
 - ✓ Herramienta de desarrollo ECL
- ✓ ECL CLI
 - ✓ Interface de línea de comandos
- ✓ ECL Watch
 - ✓ Herramienta web de gestión y supervisión

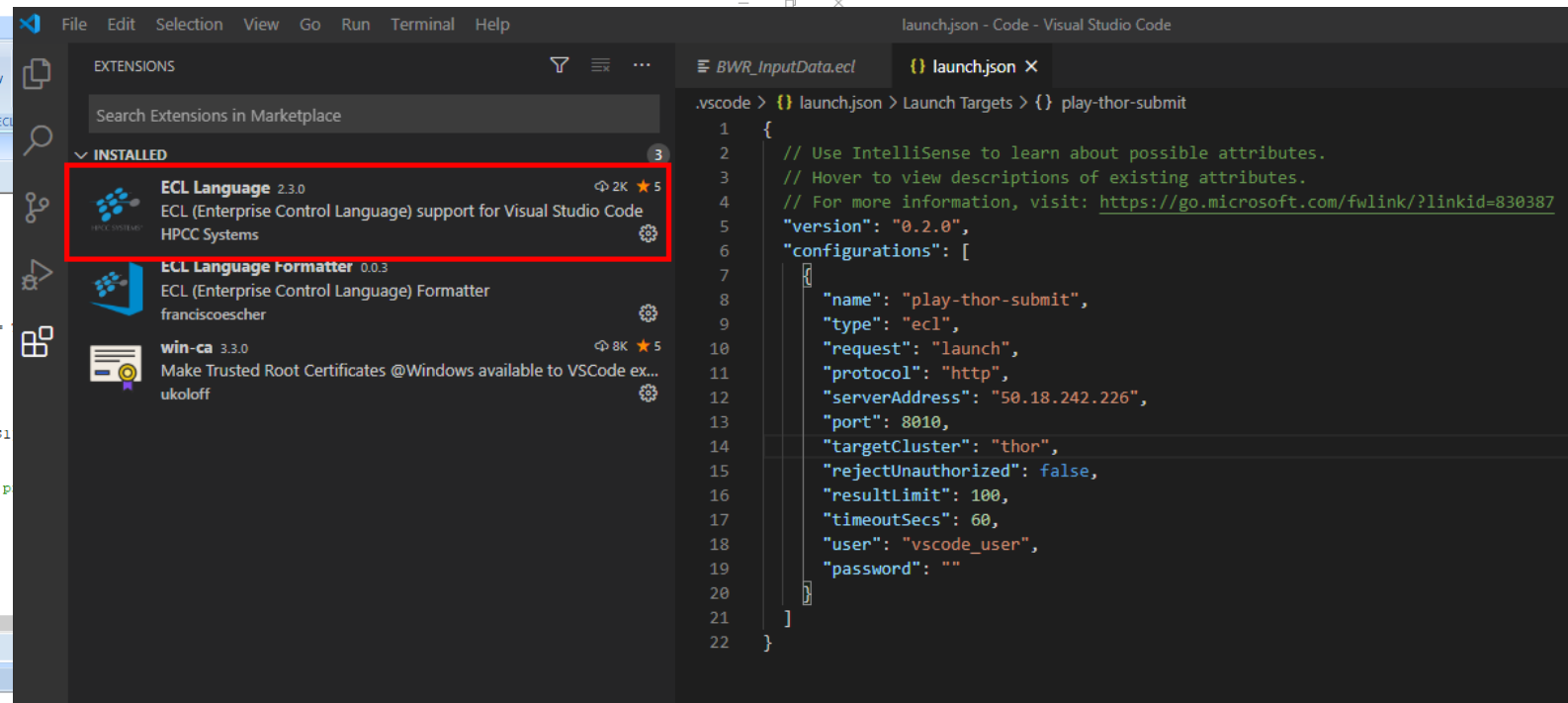


ECL IDE

✓ IDE (Win)

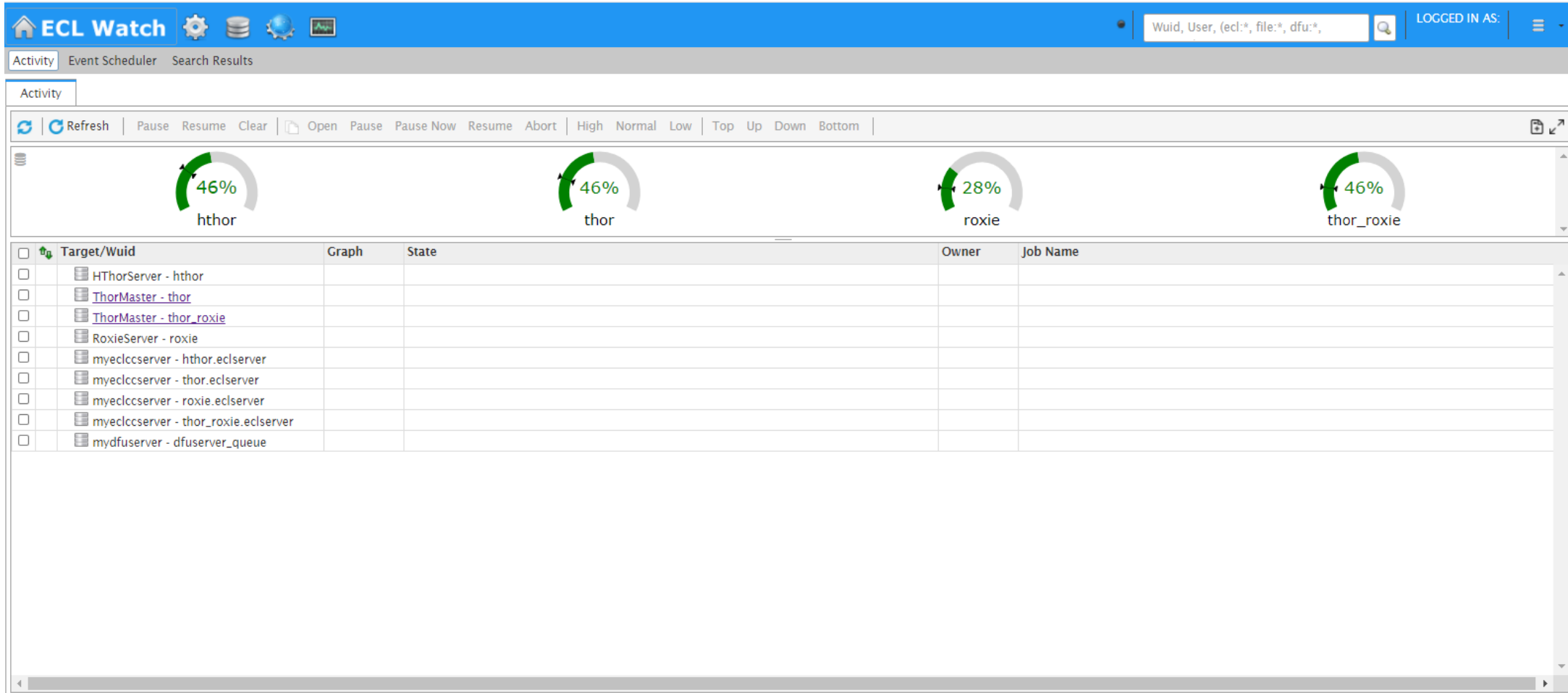


✓ VSCode (Ux/MacOS)



ECL Watch

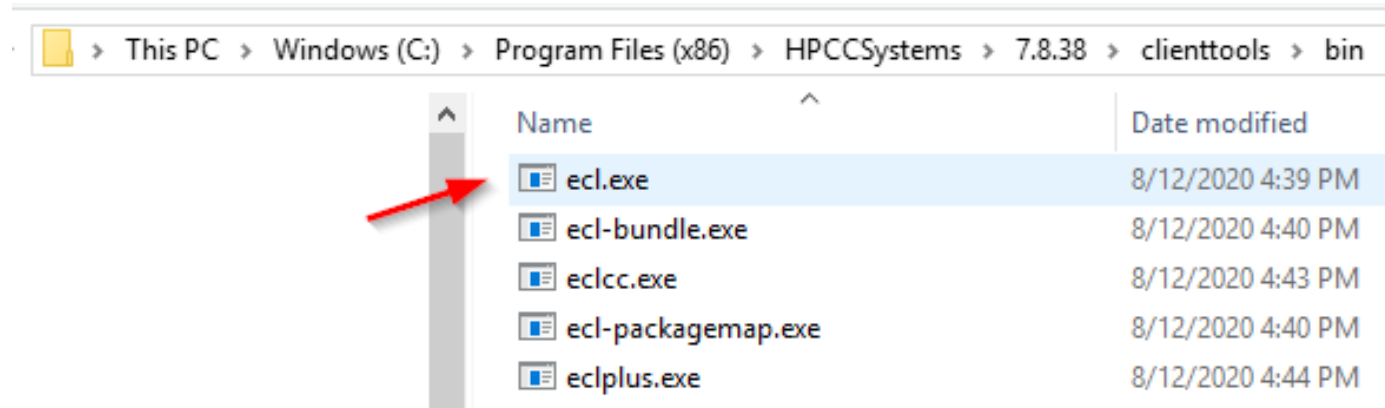
✓ Interface web (<ip>:8010)



ECL CLI

✓ Clienttools (Win/Unix)

```
hugo@hugo-VirtualBox:/opt/HPCCSystems/bin$ ls -alrt ecl*  
-rwxr-xr-x 1 root root 26776 set 23 14:43 eclscheduler  
-rwxr-xr-x 1 root root 4016848 set 23 14:43 ecl-roxie  
-rwxr-xr-x 1 root root 5970920 set 23 14:43 ecl-queries  
-rwxr-xr-x 1 root root 5958640 set 23 14:43 eclplus  
-rwxr-xr-x 1 root root 1169568 set 23 14:43 ecl-packagemap  
-rwxr-xr-x 1 root root 67736 set 23 14:43 eclccserver  
-rwxr-xr-x 1 root root 236720 set 23 14:43 eclcc  
-rwxr-xr-x 1 root root 1449272 set 23 14:43 ecl-bundle  
-rwxr-xr-x 1 root root 6142992 set 23 14:43 ecl  
hugo@hugo-VirtualBox:/opt/HPCCSystems/bin$
```

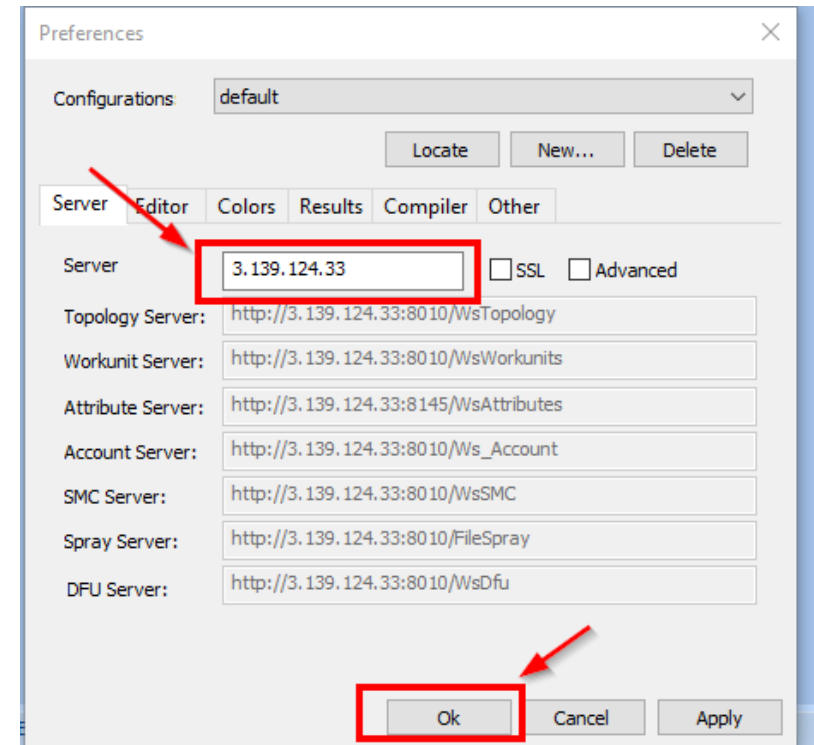
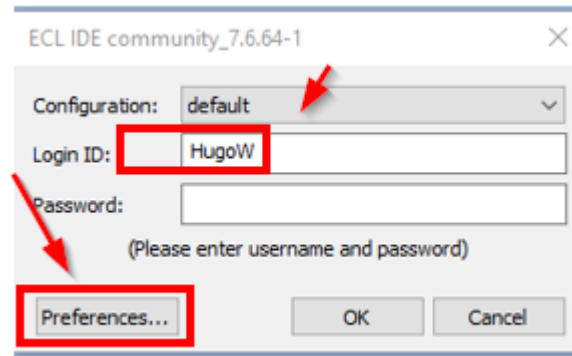
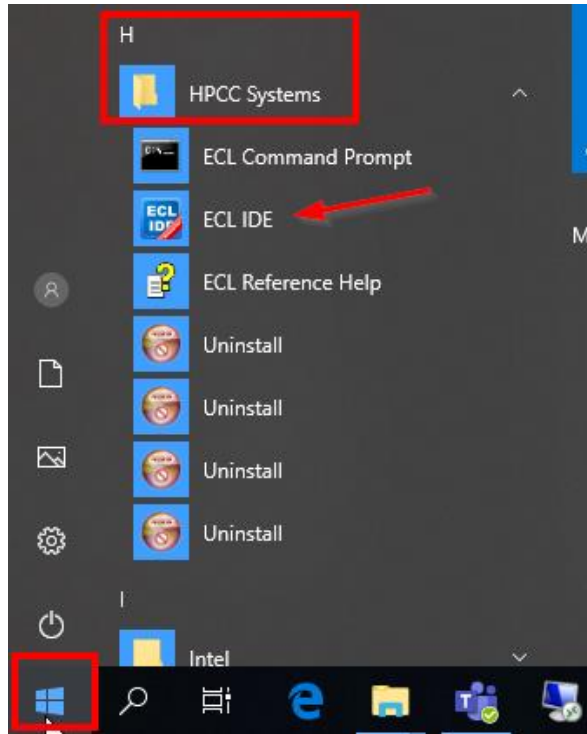


Fin de la partie 1!

Tutorial: ETL con ECL

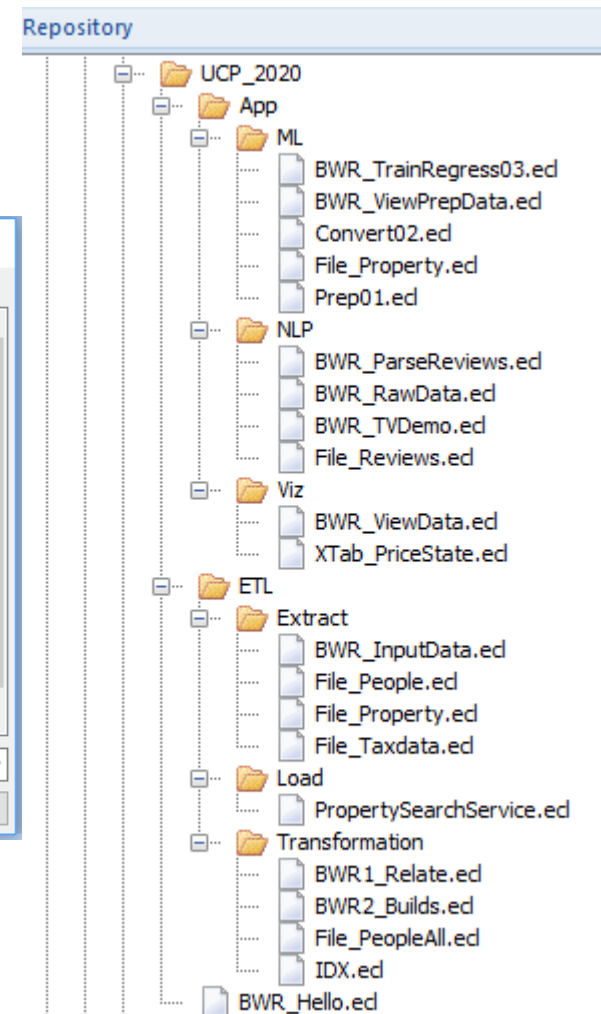
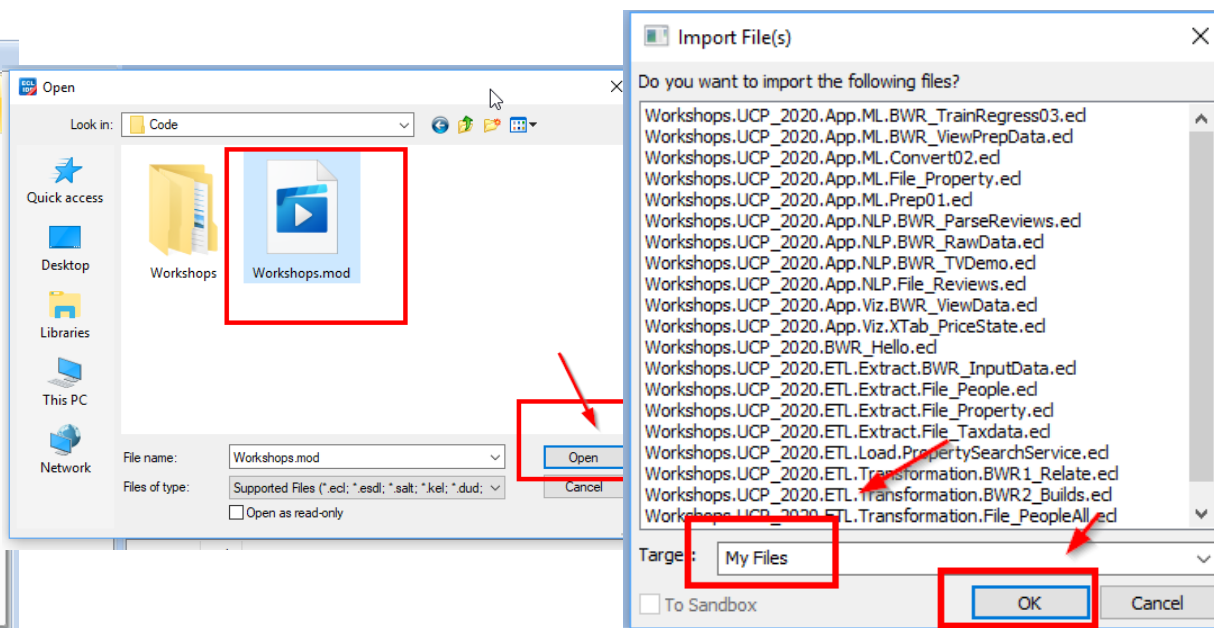
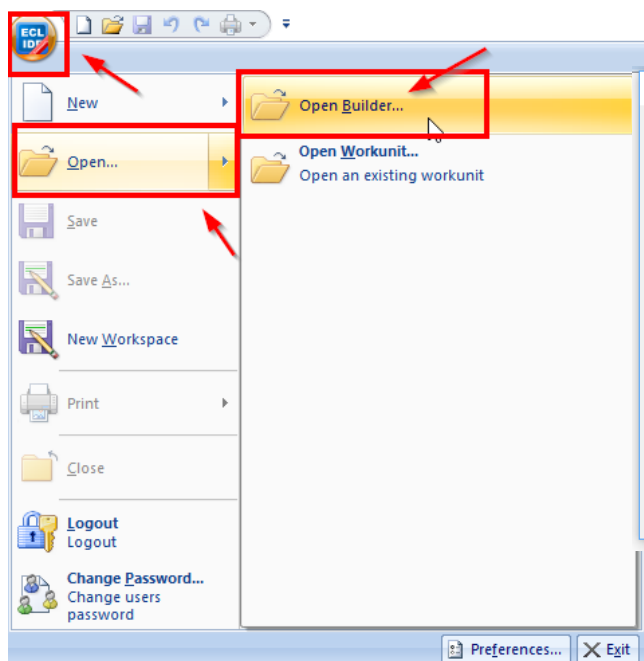
Preparación del entorno

- Clúster: <http://3.139.124.33:8010/>
- ECL IDE:

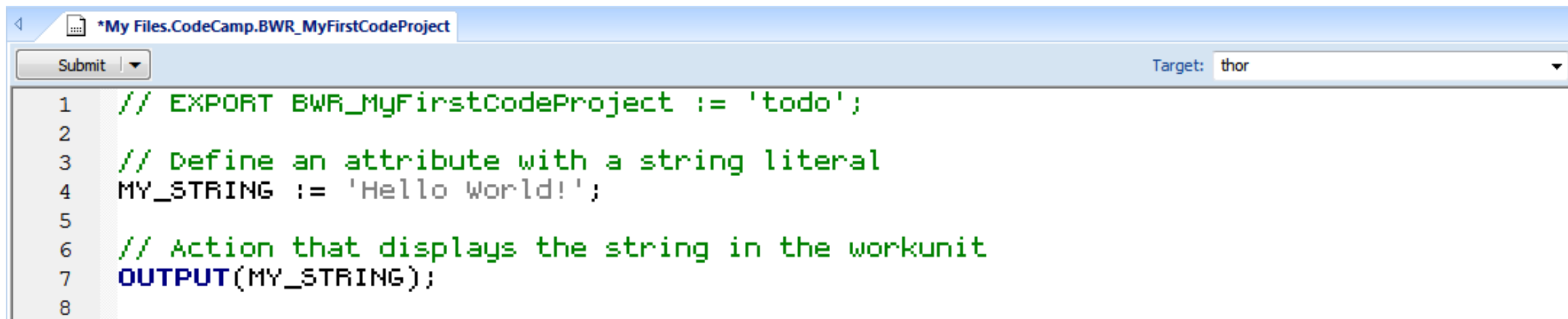


Preparación del entorno (cont.)

- Workshops.mod



ECL: Hello World!



The screenshot shows a code editor window titled '*My Files.CodeCamp.BWR_MyFirstCodeProject'. The editor has a 'Submit' button on the left and a 'Target: thor' dropdown on the right. The code is written in ECL and consists of the following lines:

```
1 // EXPORT BWR_MyFirstCodeProject := 'todo';  
2  
3 // Define an attribute with a string literal  
4 MY_STRING := 'Hello World!';  
5  
6 // Action that displays the string in the workunit  
7 OUTPUT(MY_STRING);  
8
```

Enterprise Control Language (ECL)

- Lenguaje de programación centrado en datos (dataflow)
 - Declarativa y no procesal
 - Códigos reutilizables y más pequeños
 - Biblioteca para manipulación de datos
- Compilador
 - Genera código optimizado (C++)
 - Lógica para procesamiento paralelo y distribuido

Como hacer  vs.  Que hacer

Conceptos básicos de ECL

- Estructura básica : **Nombre := Expresión;**
- ECL no distingue entre mayúsculas y minúsculas
- Se ignoran los espacios en blanco para una mejor lectura
- Comentarios en línea (//) y en bloque (/* y */)
- ECL utiliza la sintaxis de objetos

Dataset.Campo // hace referencia a un campo en un conjunto de datos

NombreDirectorio.Definicion // hace referencia a una definición en otro directorio

Objetivo del tutorial

- Servicio de consulta inmobiliaria:

propertysearchservice-manager-roxie.1 Response

Dataset: Result 1

| | lastname | firstname | id | middlename | namesuffix | filedate | gender | birthdate | propcount | | | | | | | | | | | | |
|---|----------|------------|----------------------|------------|------------|----------|--------|-----------|-----------|----------------------|------------|--------------|---------------------|--------|----------|------------|---------|--------|-------------|-------|-------|
| | | | | | | | | | | personid | propertyid | house number | house number suffix | predir | street | streettype | postdir | apt | city | state | zip |
| 1 | TAYLOR | ANJILLETTE | 17400512667362477405 | W | | 19980619 | M | 19710614 | 4 | 17400512667362477405 | 1965045 | 830 | | | SKYSAIL | AVE | | | WARRENSBURG | NY | 12885 |
| | | | | | | | | | | 17400512667362477405 | 1975231 | 838 | | NW | 000081ST | AVE | | 000002 | ST LOUIS | MO | 63131 |

Materia prima

| ## | id | firstname | lastname | middlename | namesuffix | filedate | gender | birthdate |
|----|------------------|-----------|-------------|------------|------------|----------|--------|-----------|
| 1 | 187522928604396 | PETRONICA | SPOCK | | | 20030425 | F | 19290205 |
| 2 | 214582956185891 | KIHM | DEMIRTAS | W | | 19860711 | F | 19330610 |
| 3 | 345438575926606 | DELYNN | MALSCH | T | | 20000311 | M | 19700426 |
| 4 | 562092156665191 | FOLAKE | KOSTMAN | G | | 20070922 | M | 19681006 |
| 5 | 599574955213581 | ORA | HUBERT | | | 20111011 | M | |
| 6 | 630037699819979 | KUOR | LUHCS | | | 20100402 | M | |
| 7 | 638971319693497 | ADEREMI | HOWD | | | 20000422 | M | |
| 8 | 1028541850646460 | IRA | DUNHAMPEARS | | | 20130512 | F | 19861204 |
| 9 | 1096143903819059 | TAMASINE | LUECKE | G | | 20071229 | M | |
| 10 | 1151459511906416 | SHARNAE | LITINAS | E | | 19981017 | M | 19690104 |

| ## | personid | propertyid | house_number | house_number_suffix | predir | street | streettype | postdir | apt | city | state | zip |
|----|-----------------|------------|--------------|---------------------|--------|-------------|------------|---------|--------|-------------------|-------|------|
| 1 | 187522928604396 | 828195 | 144 | | | MCKIERNAN | DR | | | WALNUT CREEK | CA | 9459 |
| 2 | 187522928604396 | 1144455 | 281 | | | CENTER | ST | | | BALTIMORE | MD | 2113 |
| 3 | 187522928604396 | 1494347 | 483 | | | NEWTON | RD | | | FLAGSTAFF | AZ | 8601 |
| 4 | 187522928604396 | 1910847 | 802 | | | HATCHERY | CT | | | WOODLAND | WA | 9867 |
| 5 | 187522928604396 | 4267562 | 5007 | | E | ROY ROGERS | RD | | | TROY | MI | 4808 |
| 6 | 187522928604396 | 4888602 | 7607 | | | PEBBLESTONE | DR | | 000009 | KERNVILLE | CA | 9323 |
| 7 | 214582956185891 | 54135 | 4 | | | WAINWRIGHT | DR | | | NORTH FORT MYERS | FL | 3391 |
| 8 | 214582956185891 | 762012 | 125 | | | SHIPYARD | DR | | 000150 | MELBOURNE VILLAGE | FL | 3290 |
| 9 | 214582956185891 | 2331721 | 1190 | | | LITTLEOAK | DR | | | HOUSTON | TX | 7701 |
| 10 | 214582956185891 | 3276109 | 2506 | | | MEADOW | DR | | | LA QUINTA | CA | 9225 |

➤ People (~280 k)

➤ Property (~1.7 Mi)

➤ Taxdata (~6.9 Mi)

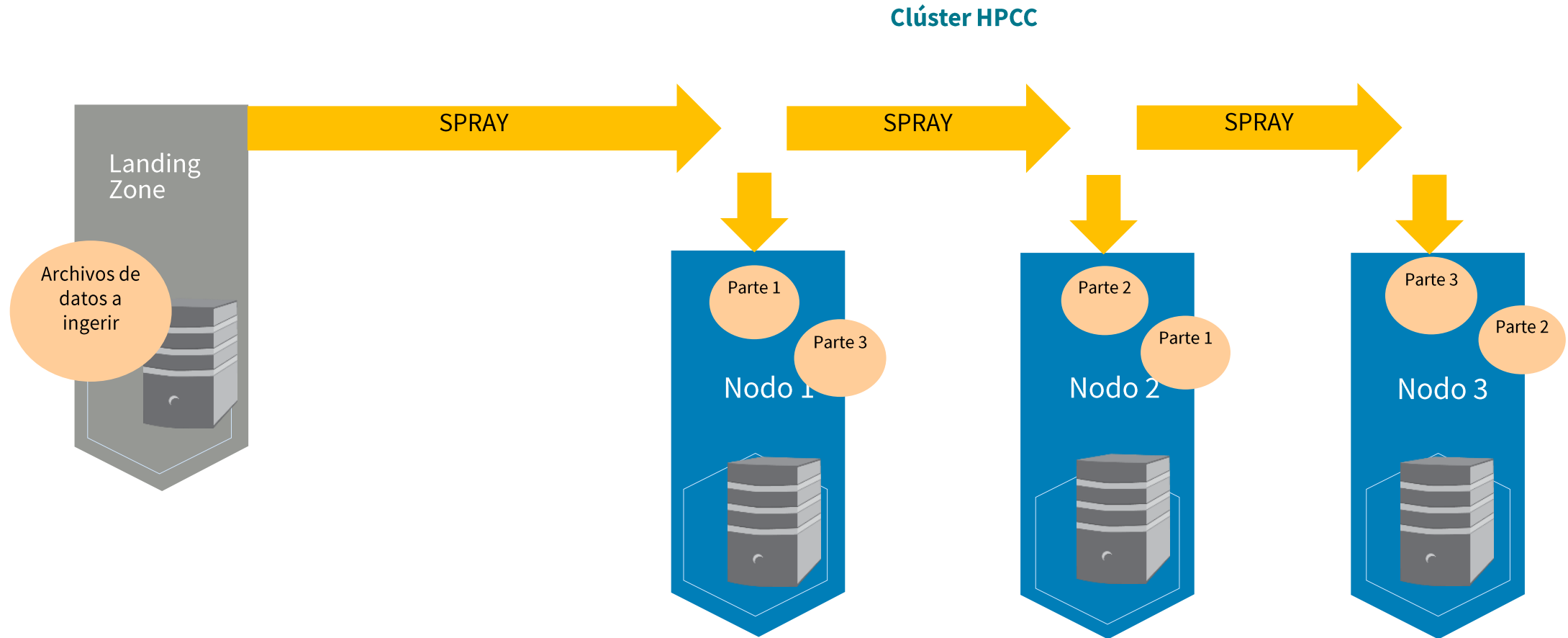
| ## | propertyid | document_year | total_val_calc | land_val_calc | improvement_value_calc | assd_total_val | tax_amount | mkt_total_val | mkt_land_val | mkt_improvement_val |
|----|------------|---------------|----------------|---------------|------------------------|----------------|------------|---------------|--------------|---------------------|
| 1 | 1 | 0000 | 101400 | 17600 | 83800 | 101400 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0000 | 107600 | 17600 | 90000 | 107600 | 0 | 0 | 0 | 0 |
| 3 | 3 | 0000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 3 | 0000 | 51353 | 8259 | 43094 | 51353 | 0 | 0 | 0 | 0 |
| 5 | 3 | 2006 | 107000 | 21400 | 85600 | 10700 | 61840 | 107000 | 21400 | 85600 |
| 6 | 4 | 0000 | 1852 | 1852 | 0 | 0 | 2870 | 1852 | 1852 | 0 |
| 7 | 4 | 0000 | 1852 | 1852 | 0 | 0 | 2928 | 1852 | 1852 | 0 |
| 8 | 4 | 2004 | 50500 | 10100 | 40400 | 4895 | 59978 | 50500 | 10100 | 40400 |
| 9 | 4 | 2004 | 50500 | 10100 | 40400 | 5050 | 62154 | 50500 | 10100 | 40400 |
| 10 | 4 | 2013 | 89000 | 17800 | 71200 | 8900 | 75690 | 89000 | 17800 | 71200 |

Definición de extracción:

Importar y limpiar datos de diferentes fuentes

- Importación de datos brutos
- Definición de la estructura de datos
- Análisis del perfil de datos

Extracción: spray de datos



Las partes del archivo se referencian en ECL como un solo archivo lógico ...

Extracción: Ejecutando el spray

<http://3.139.124.33:8010/> (ECL Watch)

The screenshot shows the ECL Watch web interface. At the top, there are icons for Logical Files, Landing Zones, Workunits, and XRef. The 'Landing Zones' tab is active, showing a list of files and folders. The 'Spray' button is highlighted with a red box and an arrow. The 'people', 'property', and 'taxdata' files are selected with checkboxes, also highlighted with a red box.

| Name |
|--|
| mydropzone [/var/lib/HPCCSystems/mydropzone] |
| 10.0.0.28 |
| DENORMALIZE |
| TrainingADVECL.MOD |
| <input checked="" type="checkbox"/> people |
| <input checked="" type="checkbox"/> property |
| <input checked="" type="checkbox"/> taxdata |
| vehicle |
| JSON |
| NLP |
| SUPERFILES |
| WORKSHOP_UCP |
| XML |
| YELP |
| var |

The screenshot shows the 'Fixed' configuration window in ECL Watch. The 'Group' is 'mythor', 'Queue' is 'dfuserver_queue', and 'Target Scope' is 'CLASS::HPCC::XXX'. The 'Target Name' and 'Record Length' table is shown. The 'Options' section has 'Overwrite' checked and 'Delayed replication' checked. The 'Spray' button is highlighted with a red box.

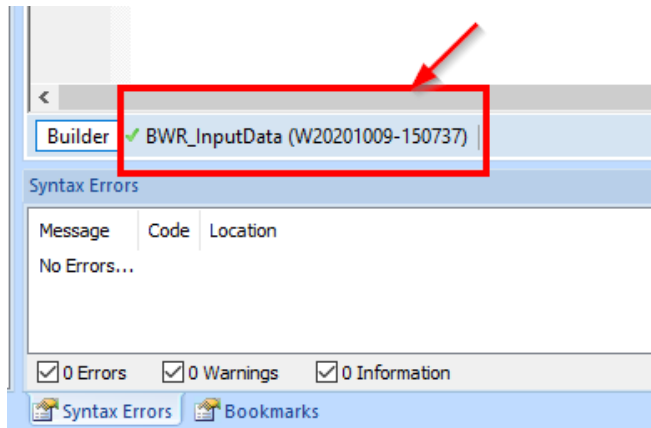
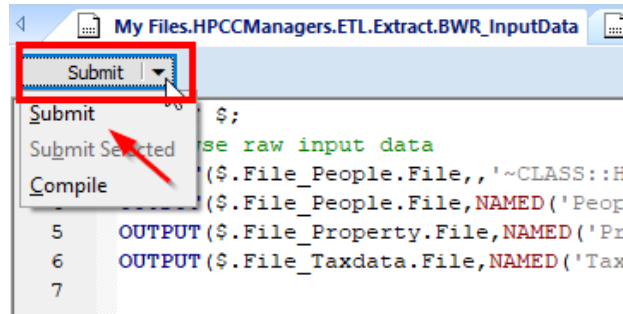
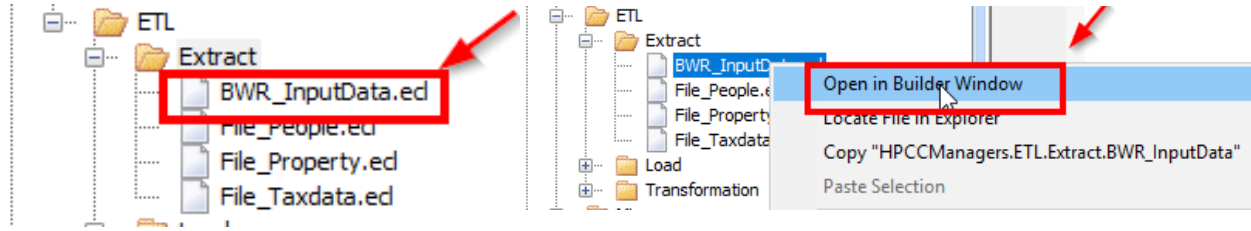
| Target Name | Record Length |
|-------------|---------------|
| people | 82 |
| property | 154 |
| taxdata | 68 |

Options:

- Overwrite: ☒
- No Split: ☐
- Compress: ☐
- Expire in (days):
- Replicate: ☐
- No Common: ☐
- Fail If No Source File: ☐
- Delayed replication: ☒

Spray

Extracción: comprobación del spray



| ## | id | firstname | lastname | middlename | namesuffix | filedate | gender | birthdate |
|----|------------------|-----------|-------------|------------|------------|----------|--------|-----------|
| 1 | 187522928604396 | PETRONICA | SPOCK | | | 20030425 | F | 19290205 |
| 2 | 214582956185891 | KIHM | DEMIRTAS | W | | 19860711 | F | 19330610 |
| 3 | 345438575926606 | DELYNN | MALSCH | T | | 20000311 | M | 19700426 |
| 4 | 562092156665191 | FOLAKE | KOSTMAN | G | | 20070922 | M | 19681006 |
| 5 | 599574955213581 | ORA | HUBERT | | | 20111011 | M | |
| 6 | 630037699819979 | KUOR | LUHCS | | | 20100402 | M | |
| 7 | 638971319693497 | ADEREMI | HOWD | | | 20000422 | M | |
| 8 | 1028541850646460 | IRA | DUNHAMPEARS | | | 20130512 | F | 19861204 |
| 9 | 1096143903819059 | TAMASINE | LUECKE | G | | 20071229 | M | |
| 10 | 1151459511906416 | SHARNAE | LITINAS | E | | 19981017 | M | 19690104 |

| ## | personid | propertyid | house_number | house_number_suffix | predir | street | streettype | postdir | apt | city | state | zip |
|----|-----------------|------------|--------------|---------------------|--------|-------------|------------|---------|--------|-------------------|-------|------|
| 1 | 187522928604396 | 828195 | 144 | | | MCKIERNAN | DR | | | WALNUT CREEK | CA | 9459 |
| 2 | 187522928604396 | 1144455 | 281 | | | CENTER | ST | | | BALTIMORE | MD | 2113 |
| 3 | 187522928604396 | 1494347 | 483 | | | NEWTON | RD | | | FLAGSTAFF | AZ | 8601 |
| 4 | 187522928604396 | 1910847 | 802 | | | HATCHERY | CT | | | WOODLAND | WA | 9867 |
| 5 | 187522928604396 | 4267562 | 5007 | | E | ROY ROGERS | RD | | | TROY | MI | 4808 |
| 6 | 187522928604396 | 4888602 | 7607 | | | PEBBLESTONE | DR | | 000009 | KERNVILLE | CA | 9323 |
| 7 | 214582956185891 | 54135 | 4 | | | WAINWRIGHT | DR | | | NORTH FORT MYERS | FL | 3391 |
| 8 | 214582956185891 | 762012 | 125 | | | SHIPYARD | DR | | 000150 | MELBOURNE VILLAGE | FL | 3290 |
| 9 | 214582956185891 | 2331721 | 1190 | | | LITTLEOAK | DR | | | HOUSTON | TX | 7701 |
| 10 | 214582956185891 | 3276109 | 2506 | | | MEADOW | DR | | | LA QUINTA | CA | 9225 |

| ## | propertyid | document_year | total_val_calc | land_val_calc | improvement_value_calc | assd_total_val | tax_amount | mkt_total_val | mkt_land_val | mkt_improvement_val |
|----|------------|---------------|----------------|---------------|------------------------|----------------|------------|---------------|--------------|---------------------|
| 1 | 1 | 0000 | 101400 | 17600 | 83800 | 101400 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0000 | 107600 | 17600 | 90000 | 107600 | 0 | 0 | 0 | 0 |
| 3 | 3 | 0000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 3 | 0000 | 51353 | 8259 | 43094 | 51353 | 0 | 0 | 0 | 0 |
| 5 | 3 | 2006 | 107000 | 21400 | 85600 | 107000 | 61840 | 107000 | 21400 | 85600 |
| 6 | 4 | 0000 | 1852 | 1852 | 0 | 0 | 2870 | 1852 | 1852 | 0 |
| 7 | 4 | 0000 | 1852 | 1852 | 0 | 0 | 2928 | 1852 | 1852 | 0 |
| 8 | 4 | 2004 | 50500 | 10100 | 40400 | 4895 | 59978 | 50500 | 10100 | 40400 |
| 9 | 4 | 2004 | 50500 | 10100 | 40400 | 5050 | 62154 | 50500 | 10100 | 40400 |
| 10 | 4 | 2013 | 89000 | 17800 | 71200 | 8900 | 75690 | 89000 | 17800 | 71200 |

Extracción: análisis de datos

Reporte de profiling:

The screenshot shows the ECL Watch application interface. At the top, there's a blue header with the ECL Watch logo and several icons (gear, database, network, and a line graph). Below the header, there's a navigation bar with tabs: 'Logical Files' (highlighted with a red box), 'Landing Zones', 'Workunits', and 'XRef'. Under the 'Logical Files' tab, there's a sub-header 'Logical Files'. Below this, there's a row of action buttons: 'Refresh', 'Open', 'Delete', 'Remote Copy', 'Copy', 'Rename', and 'Add'. A table with the header 'Logical Name' is displayed below the buttons. The table has one row with the value 'class::hpcc::xxx::peopledp'. A red box highlights the 'Logical Files' tab and the table entry, and a red arrow points to the 'Copy' button.

ECL Watch

Logical Files Landing Zones Workunits XRef

Logical Files class::hmw::peopledp

Summary Contents Data Patterns ECL DEF XML Superfiles File Parts Queries Graphs Workunit History Blooms Protected by

Report Raw Data Workunit

Refresh Analyze Target: thor Optimize Delete

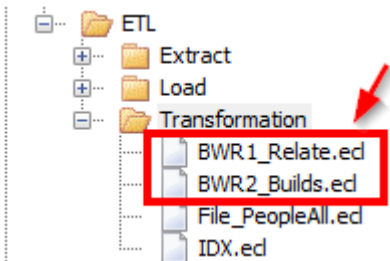
| # id | unsigned8 | Mean 9215216775812328000 Std. Deviation 5320927047810838000 | Quartiles 187522928604396 Min 4613393090456637000 25% 9205797990909421000 50% 13813182816064800000 75% 18446714708963650000 Max | Min / Max | Cardinality N/A | Popular Patterns 99999999999999999999 49% 99999999999999999999 46% 99999999999999999999 5% 99999999999999999999 0% 99999999999999999999 0% 99999999999999999999 0% |
|-------------|-----------|--|---|-----------------|--|--|
| A firstname | string15 | Min Length 1 Avg Length 6 Max Length 14 | Quartile N/A | Cardinality N/A | Popular Patterns AAAAAA 24% AAAAAAA 22% AAAAA 18% AAAAAAAAA 14% AAAA 8% AAAAAAAAA 7% Other (26) ~7% | |

Definición de transformación:

Mapeo y conversión de datos para diseños de registros estandarizados

- Designación de identificadores (recid's)
- Estandarización de campo
- Normalización o desnormalización

Transformación: desnormalización

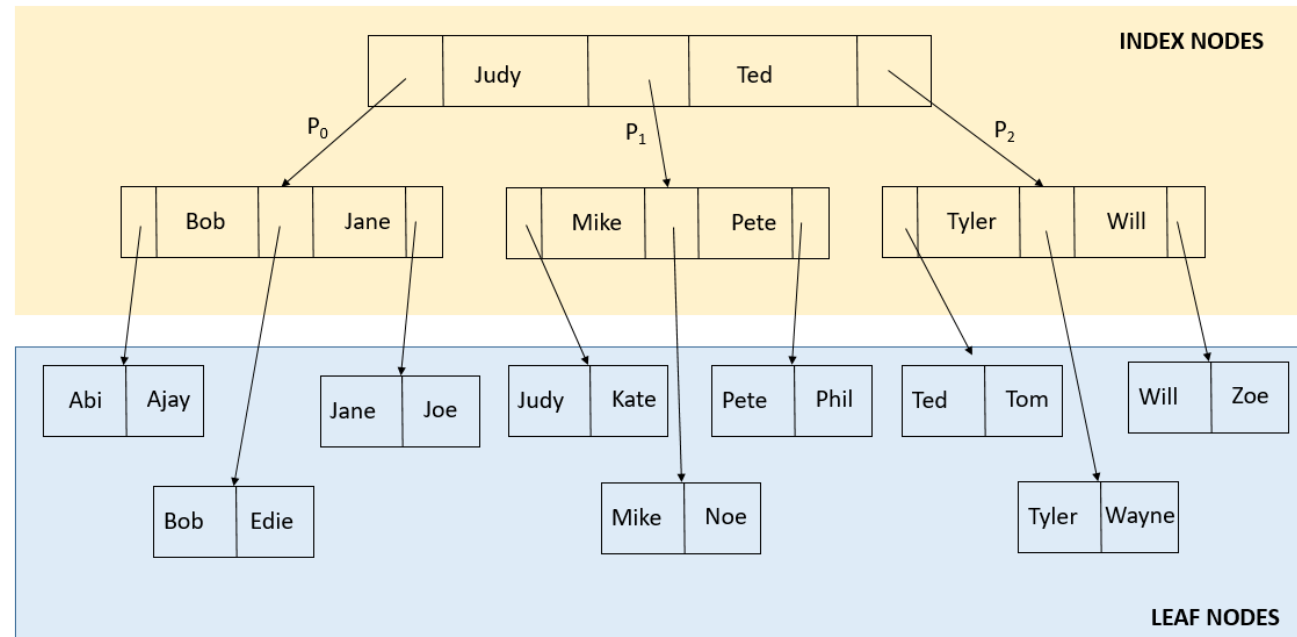


| ## | id | firstname | lastname | middlename | namesuffix | filedate | gender | birthdate | propcount | proprecs | | | | | | | | | | | |
|----|------------|-----------|----------|------------|------------|----------|--------|-----------|-----------|------------|------------|--------------|--------------|---------|-----------|------------|---------|-----|--------------|-------|-------|
| | | | | | | | | | | personid | propertyid | house_number | house_number | pre_dir | street | streettype | postdir | apt | city | state | zip |
| 1 | 1875229... | PETRONICA | SPOCK | | | 20030425 | F | 19290205 | 6 | 1875229... | 828195 | 144 | | | MCKIERNAN | DR | | | WALNUT CREEK | CA | 94597 |
| | | | | | | | | | | 1875229... | 1144455 | 281 | | | CENTER | ST | | | BALTIMORE | MD | 21136 |
| | | | | | | | | | | 1875229... | 1494347 | 483 | | | NEWTON | RD | | | FLAGSTAFF | AZ | 86011 |

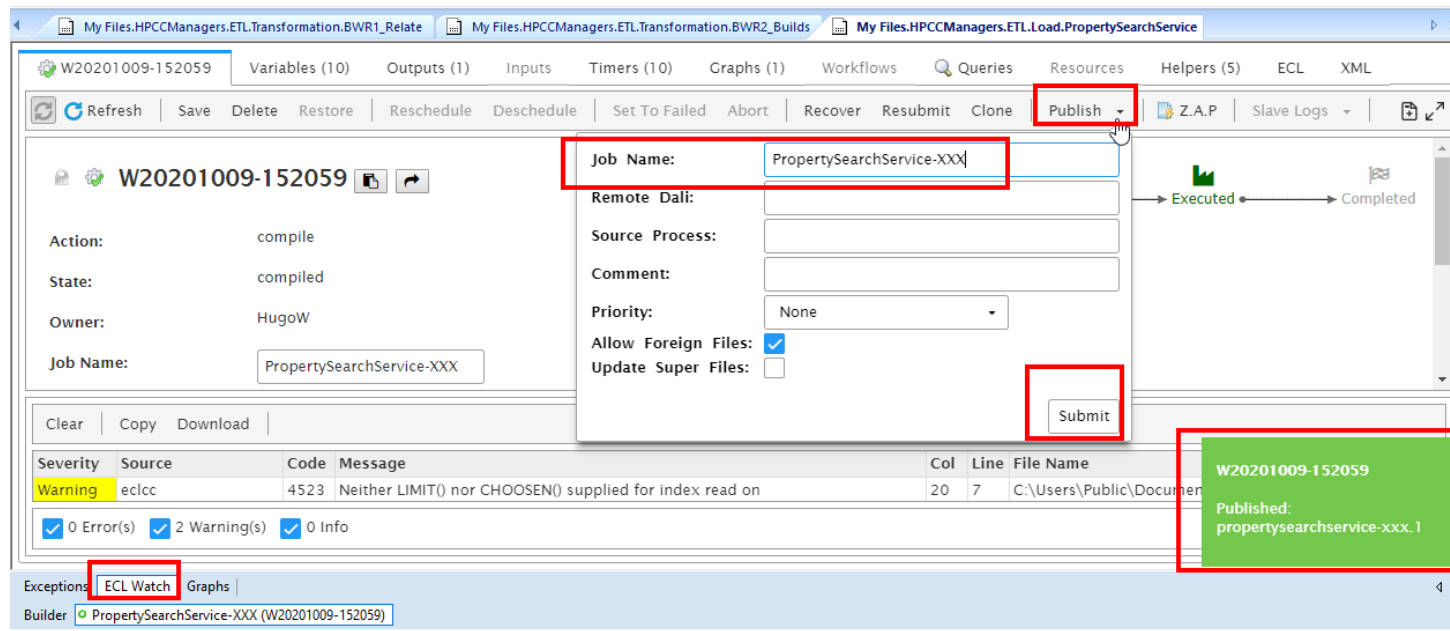
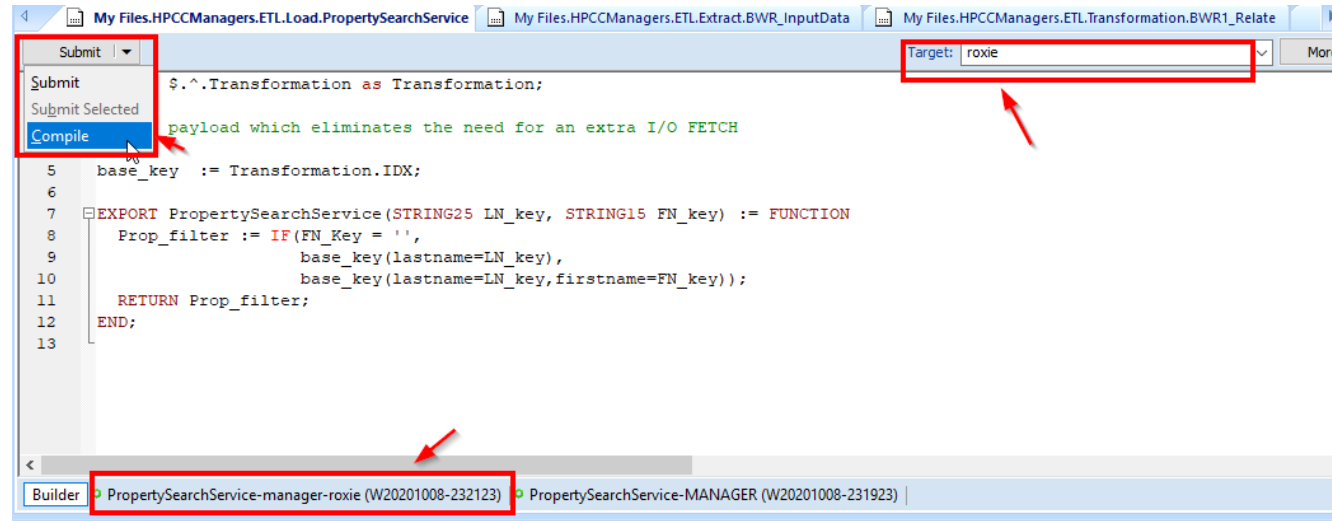
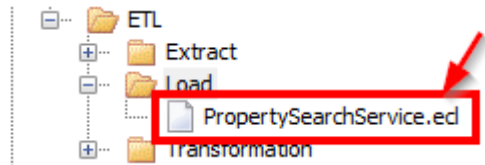
Definición de carga:

Generación de índices y disponibilidad de datos / consultas.

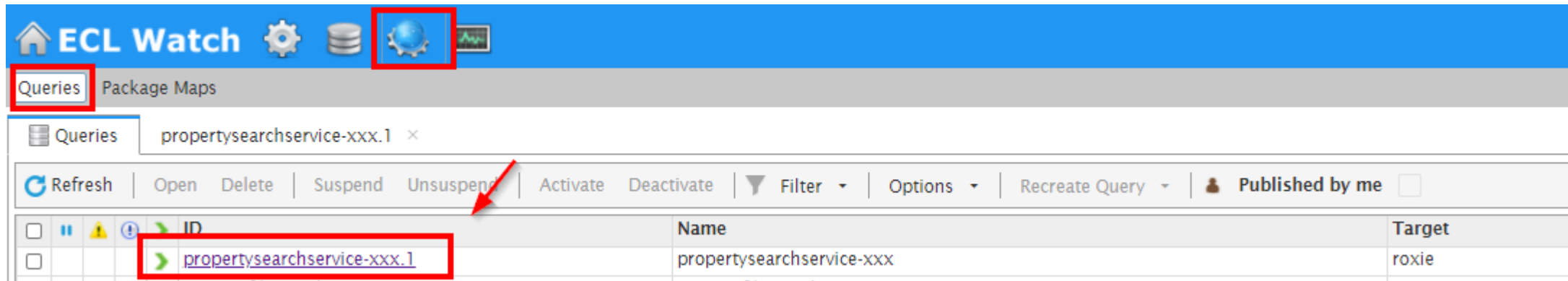
- Crear índices en el clúster THOR
- Disponibilidad de datos y consultas para un clúster ROXIE



Cargando: Publicación de la consulta

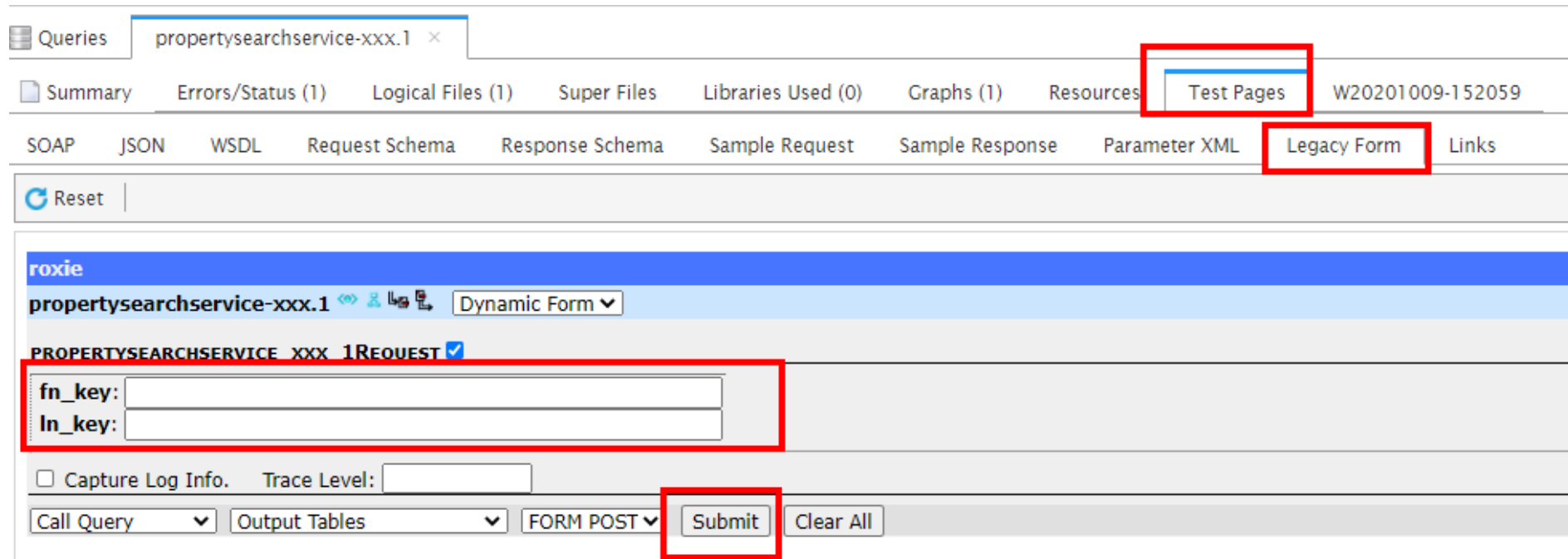


¡Servicio disponible para su uso!



The screenshot shows the ECL Watch interface. In the top navigation bar, the 'Queries' tab is highlighted. Below it, the 'propertysearchservice-xxx.1' query is selected. The query list shows the following details:

| ID | Name | Target |
|-----------------------------|---------------------------|--------|
| propertysearchservice-xxx.1 | propertysearchservice-xxx | roxie |



The screenshot shows the 'Test Pages' tab for the 'propertysearchservice-xxx.1' query. The 'Legacy Form' is displayed, and the 'Submit' button is highlighted. The form includes the following fields and controls:

- SOAP, JSON, WSDL, Request Schema, Response Schema, Sample Request, Sample Response, Parameter XML, Legacy Form, Links
- Reset
- roxie
- propertysearchservice-xxx.1
- Dynamic Form
- PROPERTYSEARCHSERVICE XXX 1REQUEST
- fn_key: [input field]
- ln_key: [input field]
- Capture Log Info. Trace Level: [input field]
- Call Query [dropdown] Output Tables [dropdown] FORM POST [dropdown] Submit [button] Clear All [button]

Fin de la partie 2!

Bundles y aplicaciones

<https://covid19.hpccsystems.com/>

THE WORLD

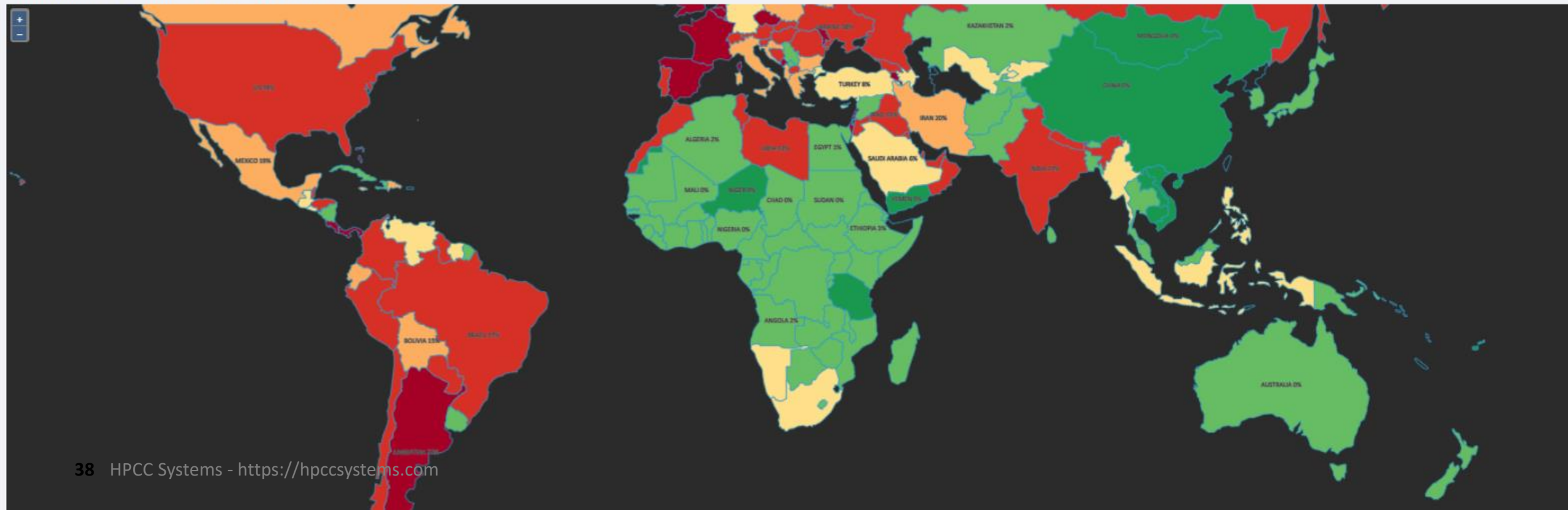
As of Oct 07, 2020, The World remains in a Stabilized state. The infection is steady ($R = 1.01$). There are currently 3,012,916 active cases. At this rate, expect to see approximately 2,129,361 new cases and 36,551 deaths per week. This is the 3rd surge in infections, which started on the week of Jun 11, 2020. With 2,129,361 new cases, this is the worst week so far for cases during this surge. The Contagion Risk is high at 18.2%. This is the likelihood of meeting an infected person during one hundred random encounters. The Case Fatality Rate (CFR) is estimated as 3.1%. Preliminary estimates suggest that 2% of the population may have been infected and are presumed immune. This is not enough to significantly slow the spread of the virus. This preliminary estimation also implies an Infection Fatality Rate (IFR) of roughly 0.6%. The Short-Term Indicator suggests that the infection is likely to worsen over the course of the next few days.

Maps

| Contagion Risk | Infection State | Weekly New Cases | Weekly New Deaths | Cases/100K | Deaths/100K | Cases | Deaths | Legend |
|----------------|-----------------|------------------|-------------------|------------|-------------|-------|--------|--------|
|----------------|-----------------|------------------|-------------------|------------|-------------|-------|--------|--------|

| Current | Historical |
|--|--|
| <p>1. Globalization</p> <p>2. Technology</p> <p>3. Demographics</p> <p>4. Environmental</p> <p>5. Political</p> <p>6. Economic</p> <p>7. Social</p> <p>8. Cultural</p> <p>9. Health</p> <p>10. Education</p> <p>11. Energy</p> <p>12. Transportation</p> <p>13. Communication</p> <p>14. Science</p> <p>15. Art</p> <p>16. Literature</p> <p>17. Music</p> <p>18. Dance</p> <p>19. Theater</p> <p>20. Cinema</p> <p>21. Television</p> <p>22. Radio</p> <p>23. Internet</p> <p>24. Mobile</p> <p>25. Cloud</p> <p>26. Big Data</p> <p>27. Artificial Intelligence</p> <p>28. Robotics</p> <p>29. Space</p> <p>30. Biotechnology</p> <p>31. Medicine</p> <p>32. Law</p> <p>33. Business</p> <p>34. Finance</p> <p>35. Marketing</p> <p>36. Advertising</p> <p>37. Public Relations</p> <p>38. Journalism</p> <p>39. Education</p> <p>40. Healthcare</p> <p>41. Environment</p> <p>42. Energy</p> <p>43. Transportation</p> <p>44. Communication</p> <p>45. Science</p> <p>46. Art</p> <p>47. Literature</p> <p>48. Music</p> <p>49. Dance</p> <p>50. Theater</p> <p>51. Cinema</p> <p>52. Television</p> <p>53. Radio</p> <p>54. Internet</p> <p>55. Mobile</p> <p>56. Cloud</p> <p>57. Big Data</p> <p>58. Artificial Intelligence</p> <p>59. Robotics</p> <p>60. Space</p> <p>61. Biotechnology</p> <p>62. Medicine</p> <p>63. Law</p> <p>64. Business</p> <p>65. Finance</p> <p>66. Marketing</p> <p>67. Advertising</p> <p>68. Public Relations</p> <p>69. Journalism</p> <p>70. Education</p> <p>71. Healthcare</p> <p>72. Environment</p> <p>73. Energy</p> <p>74. Transportation</p> <p>75. Communication</p> <p>76. Science</p> <p>77. Art</p> <p>78. Literature</p> <p>79. Music</p> <p>80. Dance</p> <p>81. Theater</p> <p>82. Cinema</p> <p>83. Television</p> <p>84. Radio</p> <p>85. Internet</p> <p>86. Mobile</p> <p>87. Cloud</p> <p>88. Big Data</p> <p>89. Artificial Intelligence</p> <p>90. Robotics</p> <p>91. Space</p> <p>92. Biotechnology</p> <p>93. Medicine</p> <p>94. Law</p> <p>95. Business</p> <p>96. Finance</p> <p>97. Marketing</p> <p>98. Advertising</p> <p>99. Public Relations</p> <p>100. Journalism</p> | <p>1. Globalization</p> <p>2. Technology</p> <p>3. Demographics</p> <p>4. Environmental</p> <p>5. Political</p> <p>6. Economic</p> <p>7. Social</p> <p>8. Cultural</p> <p>9. Health</p> <p>10. Education</p> <p>11. Energy</p> <p>12. Transportation</p> <p>13. Communication</p> <p>14. Science</p> <p>15. Art</p> <p>16. Literature</p> <p>17. Music</p> <p>18. Dance</p> <p>19. Theater</p> <p>20. Cinema</p> <p>21. Television</p> <p>22. Radio</p> <p>23. Internet</p> <p>24. Mobile</p> <p>25. Cloud</p> <p>26. Big Data</p> <p>27. Artificial Intelligence</p> <p>28. Robotics</p> <p>29. Space</p> <p>30. Biotechnology</p> <p>31. Medicine</p> <p>32. Law</p> <p>33. Business</p> <p>34. Finance</p> <p>35. Marketing</p> <p>36. Advertising</p> <p>37. Public Relations</p> <p>38. Journalism</p> <p>39. Education</p> <p>40. Healthcare</p> <p>41. Environment</p> <p>42. Energy</p> <p>43. Transportation</p> <p>44. Communication</p> <p>45. Science</p> <p>46. Art</p> <p>47. Literature</p> <p>48. Music</p> <p>49. Dance</p> <p>50. Theater</p> <p>51. Cinema</p> <p>52. Television</p> <p>53. Radio</p> <p>54. Internet</p> <p>55. Mobile</p> <p>56. Cloud</p> <p>57. Big Data</p> <p>58. Artificial Intelligence</p> <p>59. Robotics</p> <p>60. Space</p> <p>61. Biotechnology</p> <p>62. Medicine</p> <p>63. Law</p> <p>64. Business</p> <p>65. Finance</p> <p>66. Marketing</p> <p>67. Advertising</p> <p>68. Public Relations</p> <p>69. Journalism</p> <p>70. Education</p> <p>71. Healthcare</p> <p>72. Environment</p> <p>73. Energy</p> <p>74. Transportation</p> <p>75. Communication</p> <p>76. Science</p> <p>77. Art</p> <p>78. Literature</p> <p>79. Music</p> <p>80. Dance</p> <p>81. Theater</p> <p>82. Cinema</p> <p>83. Television</p> <p>84. Radio</p> <p>85. Internet</p> <p>86. Mobile</p> <p>87. Cloud</p> <p>88. Big Data</p> <p>89. Artificial Intelligence</p> <p>90. Robotics</p> <p>91. Space</p> <p>92. Biotechnology</p> <p>93. Medicine</p> <p>94. Law</p> <p>95. Business</p> <p>96. Finance</p> <p>97. Marketing</p> <p>98. Advertising</p> <p>99. Public Relations</p> <p>100. Journalism</p> |

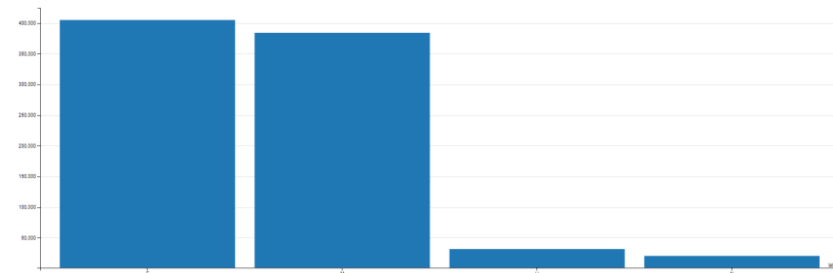
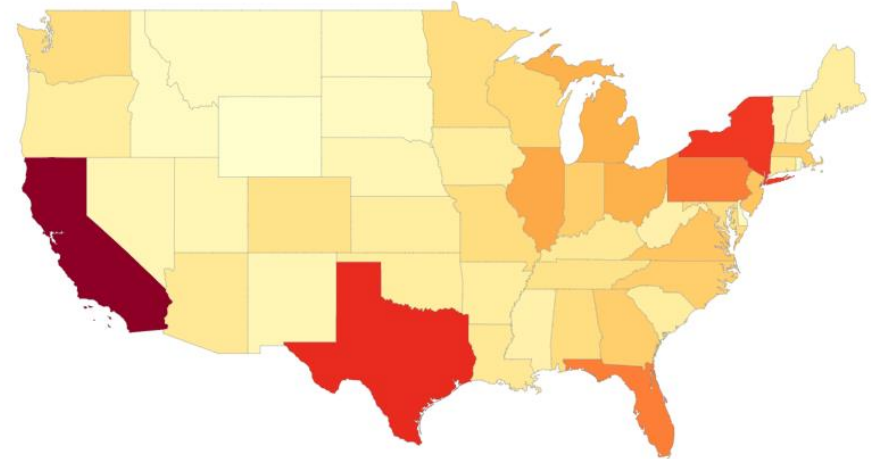
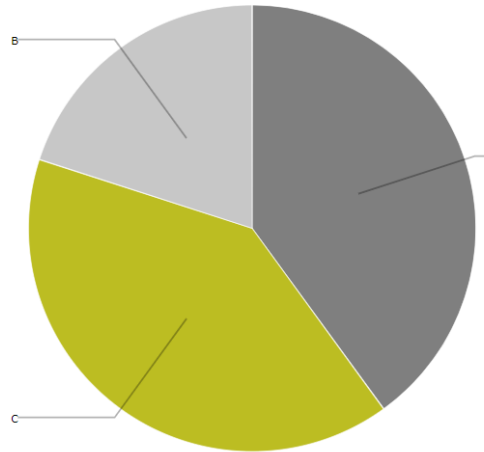
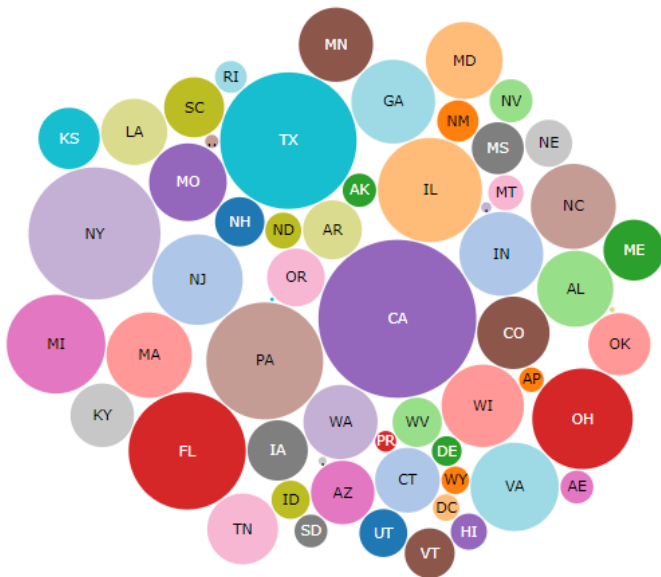
[Zoom to view more details](#) or [click on a location to view details](#).



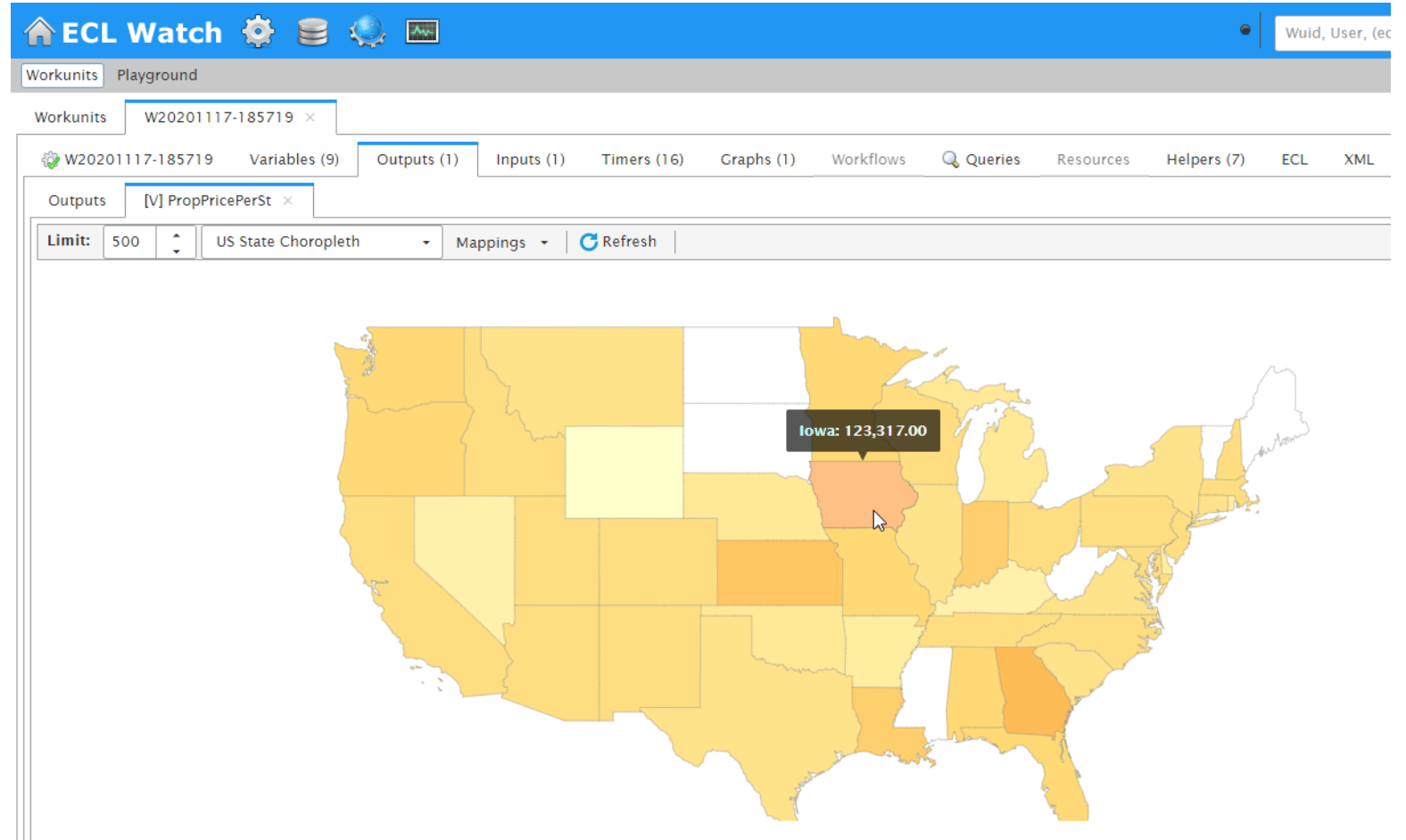
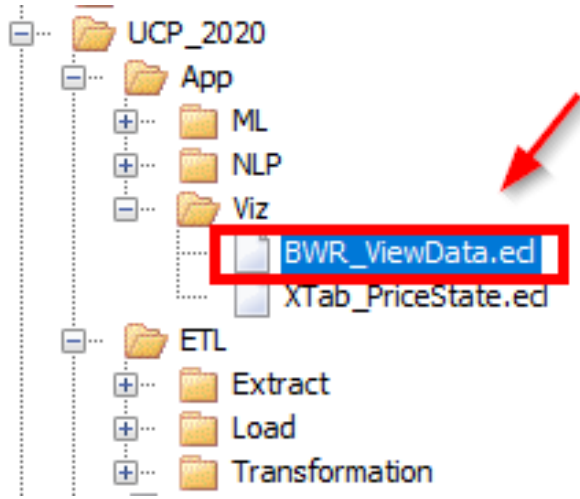
Visualización de datos

Recursos de visualización

La plataforma HPCC Systems proporciona herramientas de visualización para datos de salida a través de gráficos y mapas.



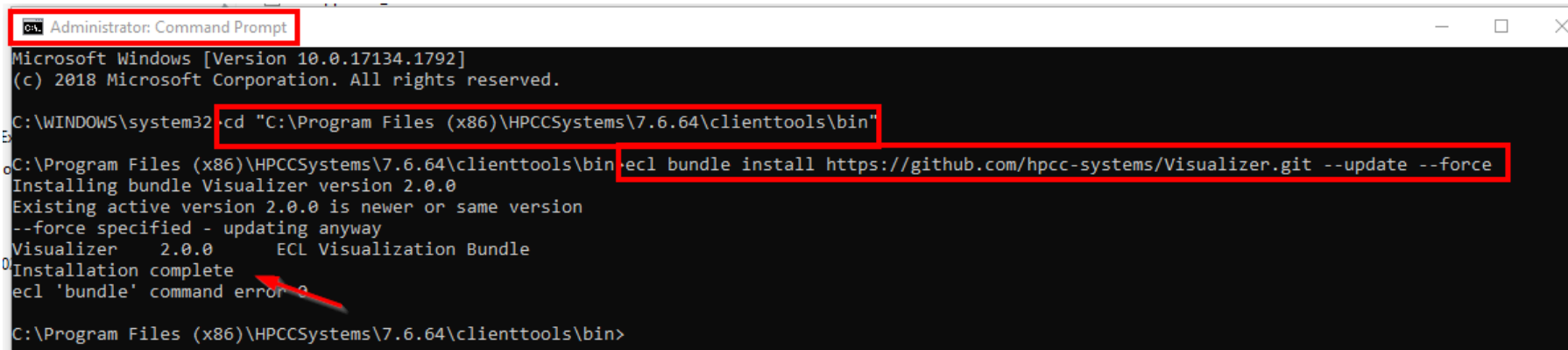
Los precios medios de propiedades



Herramientas de visualización

Los datos se pueden ver utilizando tres métodos:

- A través de la herramienta de visualización de Playground.
- A través de la pestaña "Visualizar" en la salida de cualquier unidad de trabajo.
- A través de la pestaña "Recursos" junto con el bundle de visualización ECL.

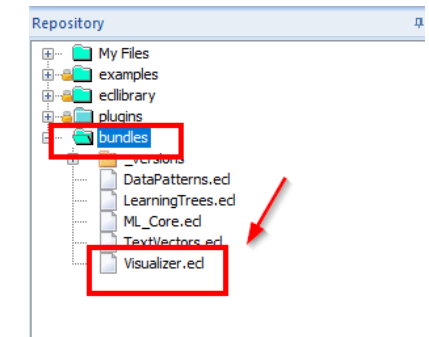


```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.17134.1792]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>cd "C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin"

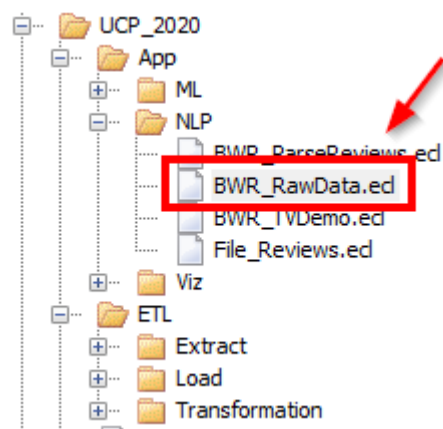
C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>ecl bundle install https://github.com/hpcc-systems/Visualizer.git --update --force
Installing bundle Visualizer version 2.0.0
Existing active version 2.0.0 is newer or same version
--force specified - updating anyway
Visualizer 2.0.0 ECL Visualization Bundle
Installation complete
ecl 'bundle' command error 0

C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>
```



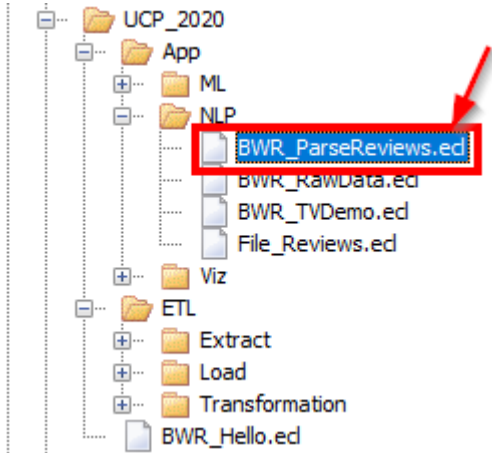
Procesamiento de Lenguaje Natural (PLN)

Los datos brutos (AirBnB)



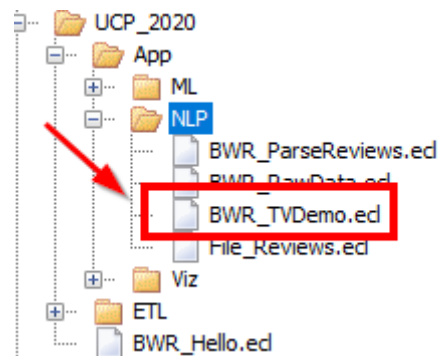
| ## | property_id | review_id | review_date | reviewer_id | reviewer_name | review_text |
|----|-------------|-----------|-------------|-------------|---------------|---|
| 1 | 7202016 | 38917982 | 2015 | 28943674 | Bianca | Cute and cozy place. Perfect location to everything! |
| 2 | 7202016 | 39087409 | 2015 | 32440555 | Frank | Kelly has a great room in a very central location. Beautiful building , architecture an |
| 3 | 7202016 | 39820030 | 2015 | 37722850 | Ian | Very spacious apartment, and in a great neighborhood. This is the kind of apartment I |
| 4 | 7202016 | 40813543 | 2015 | 33671805 | George | Close to Seattle Center and all it has to offer - ballet, theater, museum, Space Needle |
| 5 | 7202016 | 41986501 | 2015 | 34959538 | Ming | Kelly was a great host and very accommodating in a great neighborhood. She has some gre |
| 6 | 7202016 | 43979139 | 2015 | 1154501 | Barent | Kelly was great, place was great, just what I was looking for-clean, simple, well kept |
| 7 | 7202016 | 45265631 | 2015 | 37853266 | Kevin | Kelly was great! Very nice and the neighborhood and place to stay was expected and comf |
| 8 | 7202016 | 46749120 | 2015 | 24445447 | Rick | hola all bnb erz - Just left Seattle where I had a simply fantastic time for the weeken |
| 9 | 7202016 | 47783346 | 2015 | 249583 | Todd | Kelly's place is conveniently located on a quiet street in Lower Queen Anne which is an |
| 10 | 7202016 | 48388999 | 2015 | 38110731 | Tatiana | The place was really nice, clean, and the most important aspect; it was close to everyt |
| 11 | 7202016 | 49441269 | 2015 | 39852826 | Tim | The place was really nice, clean and quiet at night.Clean Linen and Towels were provide |
| 12 | 7202016 | 50490194 | 2015 | 384855 | Tony | The listing was exactly as described! Kelly's place was wonderful and cleen. it was j |
| 13 | 7202016 | 53862449 | 2015 | 21607838 | Jason | Very welcoming and a nicer place to live in the Seattle area |

Analizar los comentarios de la propiedades



| ## | prop_id | subst | verb_prep_adv | adjct |
|----|---------|---------------|----------------|-------------------------------------|
| 13 | 7202016 | Kelly | was | great |
| 14 | 7202016 | stay | was | expected and comfortable |
| 15 | 7202016 | all | was | good and mega |
| 16 | 7202016 | place | is | conveniently located |
| 17 | 7202016 | schedule | was completely | full and I |
| 18 | 7202016 | place | was really | nice |
| 19 | 7202016 | it | was close to | everything so we |
| 20 | 7202016 | place | was really | nice |
| 21 | 7202016 | Towels | were | provided and the |
| 22 | 7202016 | mattress | was | superb |
| 23 | 7202016 | Neighbourhood | is | practical with a lot of restaurants |
| 24 | 7202016 | Downtown | is | reachable by foot |
| 25 | 7202016 | Kelly | was a | fantastic host |

Extraer significado: Text Vector

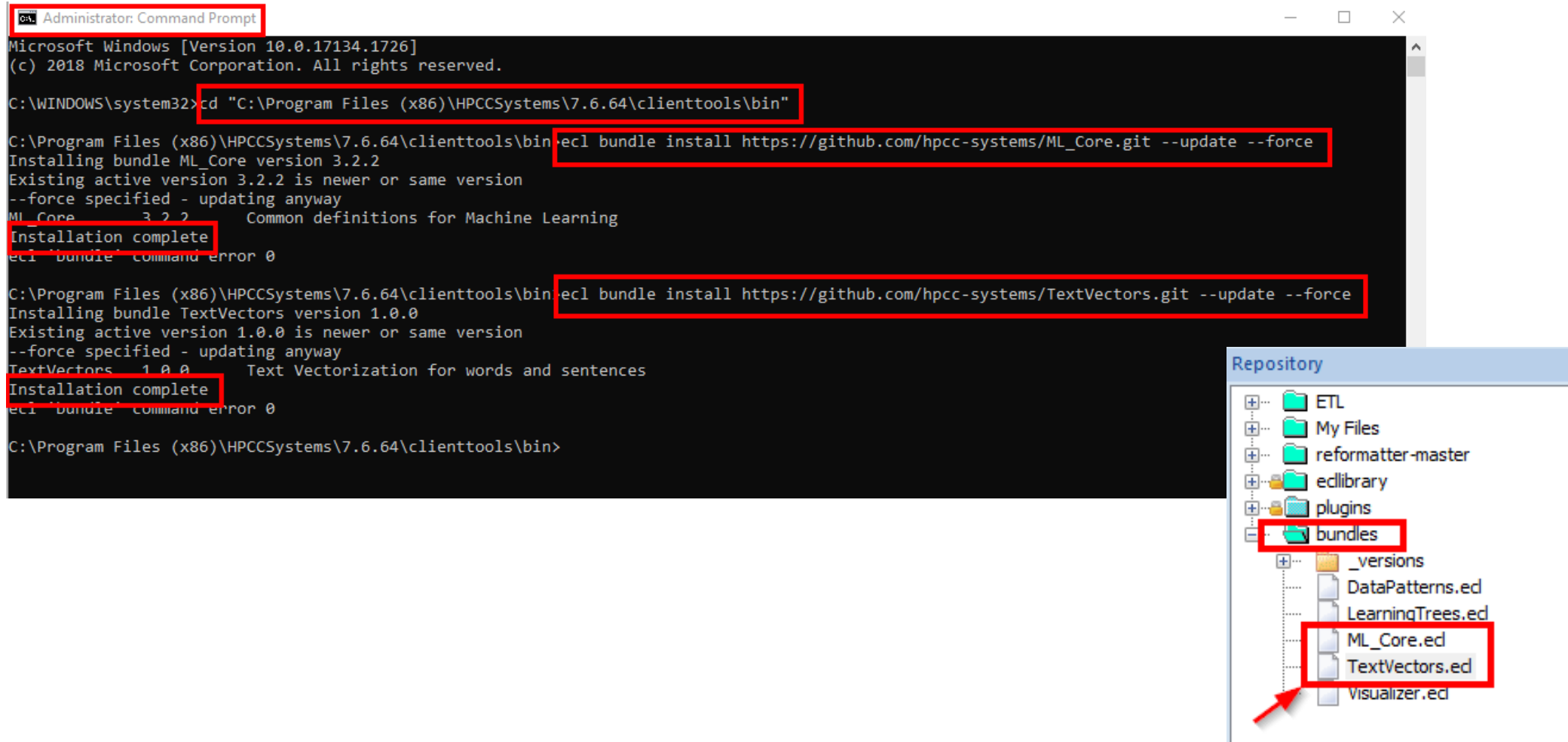


| text | closest | similarity |
|--------------------------------------|------------|---------------------|
| | Item | Item |
| location is to quiet as place is to: | quiet,stay | 1,0.999029278755188 |

| text | closest | similarity |
|---------------|--------------------------|---|
| | Item | Item |
| neighbourhood | family,awesome,beautiful | 0.9441138505935669,0.942000150680542,0.9396407008171082 |

| text | closest | similarity |
|-----------------------------|--|---------------------|
| | Item | Item |
| the apartment was spacious | Exactly as described, easy to get in and spacious.,Wonderful place! Clean, quiet, spacious, and very comfortable! Would definitely stay again! | 0.9995267987251282, |
| the neighbourhood was great | Nice quiet neighbourhood. Room was comfortable and clean.,Everything was accurate about the listing. Great location and neighbourhood. | 0.9992449283599854, |

Instalación del bundle



The image shows a Windows Command Prompt window and a file explorer window. The Command Prompt window is titled "Administrator: Command Prompt" and shows the following commands and output:

```
Microsoft Windows [Version 10.0.17134.1726]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>cd "C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin"

C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>ec1 bundle install https://github.com/hpcc-systems/ML_Core.git --update --force
Installing bundle ML_Core version 3.2.2
Existing active version 3.2.2 is newer or same version
--force specified - updating anyway
ML_Core 3.2.2 Common definitions for Machine Learning
Installation complete
ec1 bundle command error 0

C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>ec1 bundle install https://github.com/hpcc-systems/TextVectors.git --update --force
Installing bundle TextVectors version 1.0.0
Existing active version 1.0.0 is newer or same version
--force specified - updating anyway
TextVectors 1.0.0 Text Vectorization for words and sentences
Installation complete
ec1 bundle command error 0

C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>
```

The file explorer window shows the "Repository" folder with the following subfolders and files:

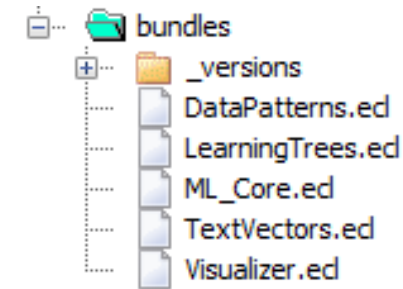
- ETL
- My Files
- reformatter-master
- edlibrary
- plugins
- bundles (highlighted with a red box)
- _versions
- DataPatterns.ed
- LearningTrees.ed
- ML_Core.ed (highlighted with a red box)
- TextVectors.ed (highlighted with a red box)
- Visualizer.ed

A red arrow points to the "bundles" folder in the file explorer.

Bundle de Machine Learning

ML en HPCC Systems

- Bundle validado, compatible y optimizado para la plataforma de sistemas HPCC (<https://hpccsystems.com/download/free-modules/machine-learning-library>)
- Proceso de instalación:
 - Fácil e independiente de la versión de la plataforma
 - *ecl bundle install* <https://github.com/hpcc-systems/<nome>.git>
- Soporte de lenguaje en R y Python:
 - Paquetes R
 - Scikit-learn
 - Keras/TensorFlow



Bundle de ML

- Base:
 - ML_Core: Machine Learning Core (https://github.com/hpcc-systems/ML_Core.git)
 - PBblas: Paralell Block Basic Linear Algebra Subsystem (<https://github.com/hpcc-systems/PBblas.git>)
- Aprendizaje supervisado
 - LinearRegression: OLS (<https://github.com/hpcc-systems/LinearRegression.git>)
 - LogisticRegression: binomial/multinomial (<https://github.com/hpcc-systems/LogisticRegression.git>)
 - GLM: General Linear Model (<https://github.com/hpcc-systems/GLM.git>)
 - SVM: Support Vector Machines (<https://github.com/hpcc-systems/SupportVectorMachines.git>)
 - LearningTrees: Árboles de decisión (<https://github.com/hpcc-systems/LearningTrees.git>)

Bundle de ML (cont.)

- Aprendizado não-supervisionado

- K-Means: agrupación de Big Data (<https://github.com/hpcc-systems/KMeans.git>)
- DBSCAN: Scalable Parallel Density-Based Spatial Clustering of Applications with Noise (<https://github.com/hpcc-systems/dbscan.git>)
- TextVectors: Vectorización de palabras, frases y oraciones (<https://github.com/hpcc-systems/TextVectors.git>)

- Aprendizaje profundo

- GNN: Generalized Neural Network (<https://github.com/hpcc-systems/GNN.git>)

Soporte a pipeline de ML



Ejemplo: Modelo de regresión

“Dado un conjunto de atributos de una propiedad (ubicación, metraje, año de construcción), ¿cómo predecir su valor?”

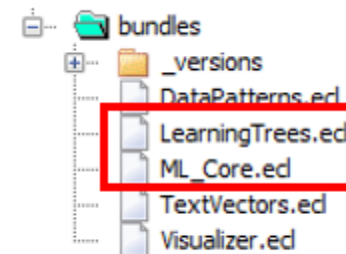
| propertyid | house_number | house_number | predir | street | street | postdir | apt | city | state | zip | total_value | assessed_value | year_acquired | land_square_foot | living_square_feet | bedrooms | full_baths |
|------------|--------------|--------------|--------|-----------------|--------|---------|--------|-----------------|-------|-------|-------------|----------------|---------------|------------------|--------------------|----------|------------|
| 828195 | 144 | | | MCKIERNAN | DR | | | WALNUT CREEK | CA | 94597 | 62614 | 62614 | 2006 | 20418 | 2485 | 3 | 2 |
| 1144455 | 281 | | | CENTER | ST | | | BALTIMORE | MD | 21136 | 105500 | 10550 | 2007 | 4807 | 1368 | 0 | 0 |
| 1494347 | 483 | | | NEWTON | RD | | | FLAGSTAFF | AZ | 86011 | 2220 | 2220 | 0 | 5654 | 1011 | 3 | 1 |
| 1910847 | 802 | | | HATCHERY | CT | | | WOODLAND | WA | 98674 | 356000 | 356000 | 0 | 6094 | 0 | 2 | 1 |
| 4267562 | 5007 | | E | ROY ROGERS | RD | | | TROY | MI | 48085 | 327253 | 327253 | 2007 | 3484 | 0 | 3 | 0 |
| 4888602 | 7607 | | | PEBBLESTONE | DR | | 000009 | KERNVILLE | CA | 93238 | 732179 | 732179 | 2010 | 19597 | 6132 | 6 | 6 |
| 48725 | 4 | | | LONG | AVE | | | SUNRISE | FL | 33323 | 271000 | 271000 | 2008 | 6880 | 2392 | 4 | 2 |
| 83528 | 6 | | | TRILLUM | LN | | | WAYLAND | MA | 02193 | 79889 | 79889 | 2007 | 7657 | 1657 | 4 | 1 |
| 94604 | 7 | | | PARMENTER | AVE | | | PLYMOUTH | MN | 55441 | 23800 | 23800 | 2005 | 19994 | 1754 | 3 | 2 |
| 220326 | 17 | | | TIMBER | RD | | | LOS ANGELES | CA | 90063 | 89000 | 89000 | 2008 | 7840 | 954 | 3 | 1 |
| 994609 | 212 | | | FREYER | DR | NE | | PHILOMONT | VA | 20131 | 59800 | 59800 | 2009 | 11199 | 1241 | 3 | 0 |
| 1836173 | 724 | | | EASTER | ST | | | ALLEN TOWN | PA | 18102 | 191600 | 191600 | 0 | 9100 | 2534 | 4 | 2 |
| 2910797 | 1903 | | | SADDLE BROOK | DR | | | CLIO | CA | 96106 | 61610 | 61610 | 2007 | 0 | 0 | 0 | 0 |
| 3083959 | 2158 | | | RIVERSIDE | DR | | | UPPER MORELA... | PA | 19006 | 90300 | 0 | 0 | 0 | 1235 | 3 | 2 |
| 3952189 | 4040 | | | GRAND VIEW | BLVD | | 000054 | RIO LINDA | CA | 95673 | 0 | 0 | 0 | 2700720 | 0 | 0 | 0 |
| 4186238 | 4726 | | | LAS PALMAS | CT | | | WAE LDER | TX | 78959 | 18816 | 18816 | 2009 | 2159 | 1320 | 0 | 0 |
| 4597143 | 6213 | | | WILSON | RD | | | ZOLFO SPRINGS | FL | 33890 | 72600 | 0 | 0 | 8496 | 0 | 3 | 1 |
| 4624905 | 6321 | | | STONEWALL | LN | | | PATERSON | NJ | 07514 | 139880 | 139880 | 2008 | 10454 | 1391 | 4 | 2 |
| 92326 | 7 | | | KNOLLCREST | DR | | | NARANJA | FL | 33032 | 76214 | 76214 | 2008 | 4800 | 930 | 2 | 0 |
| 1792852 | 704 | | | ERIN | DR | | | TRABUCO | CA | 92678 | 28010 | 28010 | 2007 | 5200 | 0 | 3 | 1 |
| 1843977 | 728 | | S | ARLINGTON HE... | RD | | | BLOOMING GRO... | TX | 76626 | 130400 | 130400 | 2007 | 36154 | 1629 | 3 | 1 |
| 4214872 | 4821 | | | MYRTLE OAK | DR | | 000025 | SAN BERNARDT | CA | 92376 | 22250 | 22250 | 2007 | 93654 | 0 | 0 | 0 |

Instalación del bundle

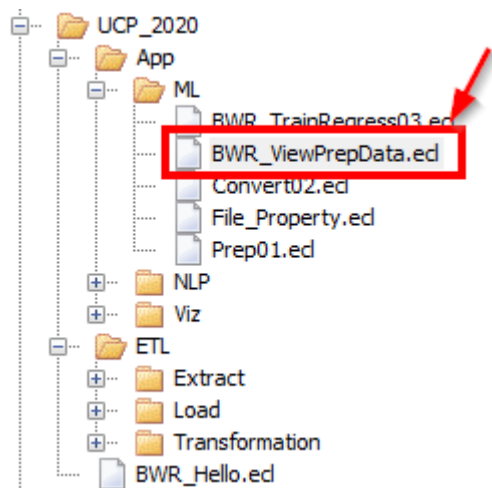
```
Command Prompt
(c) 2018 Microsoft Corporation. All rights reserved.

C:\Users\watahu01>cd "C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin"
C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>ec1 bundle install https://github.com/hpcc-systems/ML_Core.git --update --force
Installing bundle ML_Core version 3.2.2
Existing active version 3.2.2 is newer or same version
--force specified - updating anyway
ML_Core 3.2.2 Common definitions for Machine Learning
Installation complete
ec1 'bundle' command error 0

C:\Program Files (x86)\HPCCSystems\7.6.64\clienttools\bin>ec1 bundle install https://github.com/hpcc-systems/LearningTrees.git --update --force
Installing bundle LearningTrees version 1.1.1
Existing active version 1.1.1 is newer or same version
--force specified - updating anyway
LearningTrees 1.1.1 LearningTrees Bundle for Tree-based Machine Learning
Installation complete
ec1 'bundle' command error 0
```

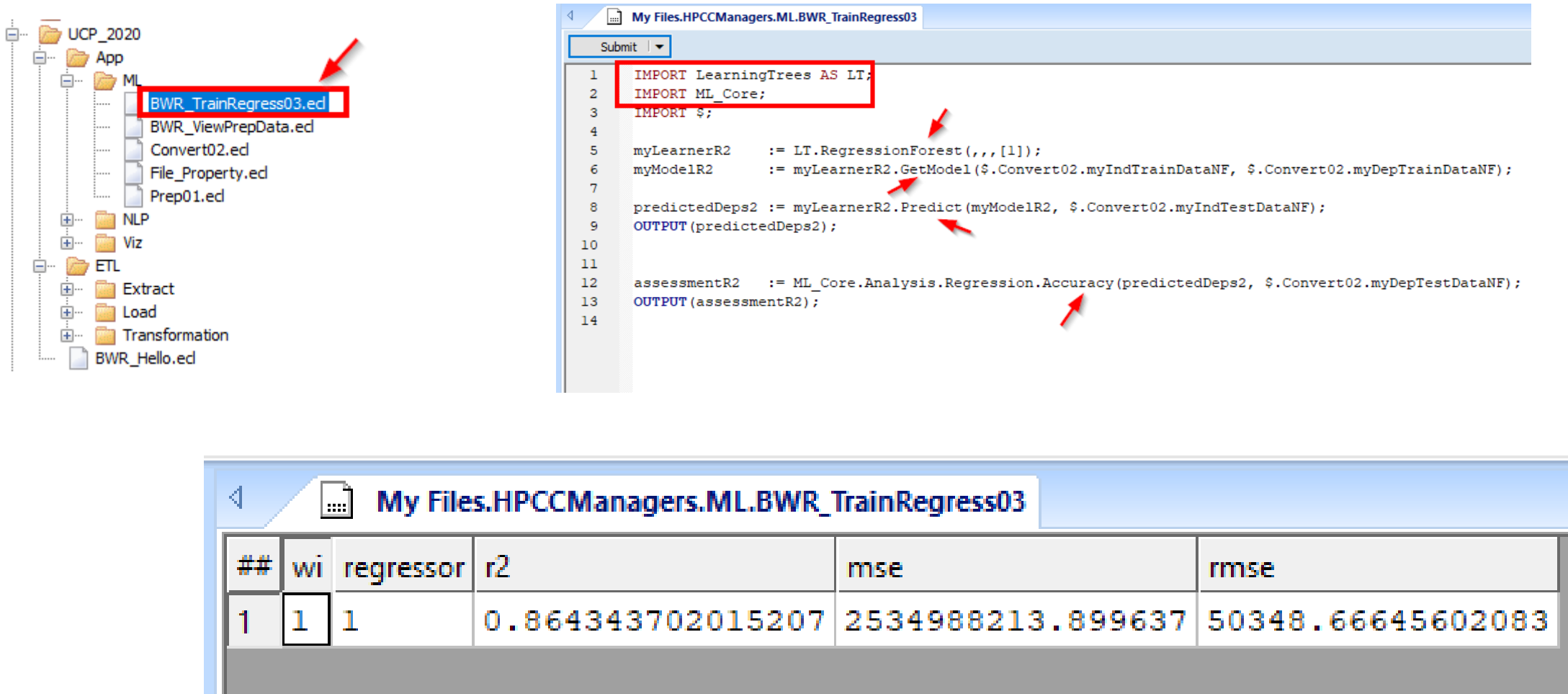


Los datos de la regresión



| ## | propertyid | zip | assessed_value | year_acquired | land_square_footage | living_square_feet | bedrooms | full_baths | half_baths | year_built | total_value | rnd |
|----|------------|-------|----------------|---------------|---------------------|--------------------|----------|------------|------------|------------|-------------|------------|
| 1 | 828195 | 94597 | 62614 | 2011 | 20418 | 2485 | 3 | 2 | 1 | 2009 | 62614 | 2681399375 |
| 2 | 4888602 | 93238 | 732179 | 2015 | 19597 | 6132 | 6 | 6 | 0 | 2010 | 732179 | 2016489933 |
| 3 | 762012 | 32904 | 96300 | 2015 | 18000 | 2357 | 4 | 2 | 1 | 2006 | 96300 | 1951014038 |
| 4 | 1565512 | 93300 | 245000 | 2012 | 5649 | 1149 | 3 | 2 | 0 | 1967 | 245000 | 2234351926 |
| 5 | 1837309 | 49333 | 9470 | 2015 | 154202 | 2284 | 3 | 2 | 0 | 1986 | 32672 | 3900215832 |
| 6 | 4542536 | 13159 | 148550 | 2015 | 7890 | 3113 | 4 | 2 | 1 | 2011 | 148550 | 1662479476 |
| 7 | 1892631 | 76136 | 840 | 2012 | 392040 | 2004 | 3 | 0 | 0 | 2003 | 188810 | 1533850982 |
| 8 | 3541423 | 95375 | 30300 | 2015 | 6534 | 2795 | 2 | 2 | 0 | 2002 | 331100 | 1750945454 |
| 9 | 3831369 | 15234 | 109163 | 2015 | 7143 | 768 | 2 | 1 | 0 | 1962 | 109163 | 3161440663 |
| 10 | 978550 | 83610 | 223900 | 2011 | 16552 | 2470 | 3 | 2 | 1 | 1999 | 223900 | 2377635704 |
| 11 | 4836980 | 7670 | 107500 | 2013 | 15681 | 1582 | 3 | 2 | 0 | 1987 | 107500 | 1313131942 |
| 12 | 769894 | 66062 | 119835 | 2013 | 6000 | 1447 | 3 | 2 | 0 | 1982 | 119835 | 3266895727 |
| 13 | 900426 | 93420 | 196639 | 2014 | 8700 | 2504 | 3 | 2 | 1 | 2008 | 196639 | 1223955664 |
| 14 | 1636831 | 23060 | 216240 | 2013 | 6540 | 1178 | 3 | 2 | 0 | 1977 | 216240 | 510070662 |
| 15 | 689976 | 78023 | 38440 | 2012 | 4420 | 1509 | 3 | 2 | 1 | 1982 | 104590 | 1855145411 |
| 16 | 3827323 | 20695 | 93600 | 2013 | 4000 | 930 | 3 | 1 | 1 | 1995 | 93600 | 2094279832 |
| 17 | 4401557 | 93637 | 25271 | 2011 | 217800 | 1916 | 3 | 2 | 1 | 2008 | 259447 | 2346788527 |
| 18 | 1502153 | 12250 | 154675 | 2016 | 1240 | 1105 | 3 | 2 | 1 | 2001 | 154675 | 2741141577 |
| 19 | 1829065 | 95403 | 44590 | 2011 | 16727 | 1896 | 4 | 2 | 1 | 2005 | 127400 | 937358568 |
| 20 | 1981287 | 34949 | 202953 | 2016 | 7700 | 1452 | 3 | 2 | 0 | 1977 | 202953 | 796231433 |

La calidad del modelo



The screenshot displays the HPCC Systems interface. On the left, a file explorer shows a project structure with folders like UCP_2020, App, ML, NLP, Viz, ETL, and sub-folders like Extract, Load, and Transformation. A file named 'BWR_TrainRegress03.ed' is highlighted with a red box and an arrow. The main area shows a code editor for 'My Files.HPCCManagers.ML.BWR_TrainRegress03' with a 'Submit' button. The code is as follows:

```
1  IMPORT LearningTrees AS LT;
2  IMPORT ML_Core;
3  IMPORT $;
4
5  myLearnerR2 := LT.RegressionForest(,,,[1]);
6  myModelR2   := myLearnerR2.GetModel($.Convert02.myIndTrainDataNF, $.Convert02.myDepTrainDataNF);
7
8  predictedDeps2 := myLearnerR2.Predict(myModelR2, $.Convert02.myIndTestDataNF);
9  OUTPUT(predictedDeps2);
10
11
12  assessmentR2 := ML_Core.Analysis.Reggression.Accuracy(predictedDeps2, $.Convert02.myDepTestDataNF);
13  OUTPUT(assessmentR2);
14
```

Below the code editor, a results table is shown with the following data:

| ## | wi | regressor | r2 | mse | rmse |
|----|----|-----------|-------------------|-------------------|-------------------|
| 1 | 1 | 1 | 0.864343702015207 | 2534988213.899637 | 50348.66645602083 |

Fin de la partie 3!

Resumen del taller

- ✓ Descripción general de la plataforma HPCC Systems
 - ✓ Definición
 - ✓ Histórico
 - ✓ Componentes
- ✓ Familiaridad con el proceso ETL en HPCC
 - ✓ Extracción
 - ✓ Transformación
 - ✓ Carga
- ✓ Comprender las aplicaciones
 - ✓ Visualización
 - ✓ PLN
 - ✓ ML

Entrenamiento online: learn.lexisnexis.com/hpcc

- Introducción a ECL (parte 1)
 - Conceptos y consultas
- Introducción a ECL (parte 2)
 - ETL con ECL
- ECL avanzado (parte 1)
 - Datos relacionales
- ECL avanzado (parte 2)
 - Superarchivos, XML / JSON y PLN
- ECL aplicado
 - Generación y automatización de código ECL
- ROXIE ECL (parte 1)
 - Índices y consultas
- ROXIE ECL (parte 2)
 - Optimización de consultas
- Machine Learning con HPCC Systems
 - Fundamentación para uso de los bundles
- Administración de sistemas
 - Conceptos básicos y funcionamiento
- HPCC para gerentes
 - Descripción general y aplicaciones de la plataforma

Enlaces útiles

- Sitio principal : hpccsystems.com/es
- Primeros pasos : hpccsystems.com/es/about
- Download: hpccsystems.com/es/download
- Foro de la Comunidad: hpccsystems.com/forums



Unete a la comunidad

Regístrese en hpccsystems.com