API Pricing Recalculate

Dataiku API Service: pricing

Created in API designer menu in dataiku project.

Endpoint: Recalculate

Background

The Pricing backend initially finds similarities between CPCs and from the 10 closest or so, compute a recommended price for the target CPC (median + business rules, overlays, etc.)

However some of the initial neighbors CPCs might not be comparable from business point of view and they remove them. The new recommended price must be recomputed, that is where the API is used.

Goal

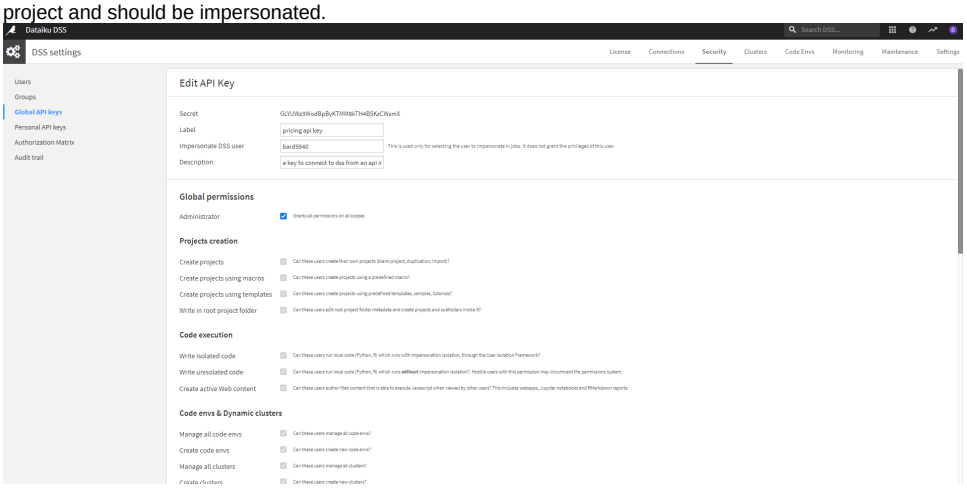
Recalculate the recommended price given the target and a list of neighbor CPCs.

Requirements

- Code shared between API and project, not duplicated
- Access to project's intermediate datasets to re compute only what is necessary
- Dataiku project API Webapp connections in different environment should match

Solution

- Code shared between project initial computation and API is in functions in project's library, not directly in recipes
- Datasets: a mix of PostgreSQL tables and managed folders in project, so we choose to connect to project's and access datasets with dataiku api (alternative would have been direct connection to postgresql if only postgresql tables for example, etc)
- API connects to remote DSS instance using a "Global API key" that should be set at instance level. This key must have appropriate rights for project and should be impersonated.



- Remote DSS is set depending on API host

```
# Host configuration
# hostname is something like: dku-mad-pricing-on-dss-test-standard-1-855fdd5784-kqvcx
hostname = socket.gethostname()
if hostname == "ce-bda-dss-design-ew1-prod":  # PROD DESIGN
    dataiku.set_remote_dss("http://ce-bda-dss-design-ew1-prod:10000", "GLYUWz3WodBpByKTMM8kTH4BSKzCWsmX")
    os.environ['DKU_CURRENT_PROJECT_KEY'] = 'PBE_MASTER'
    version = get_config_version()
elif "dss-prod-standard" in hostname:  # PROD AUTOMATION
    dataiku.set_remote_dss("http://ce-bda-dss-automation-ew1-prod:10000", "3rWX9t9OtkOCflOpBACo5jMrpQRhnwsN")
    os.environ['DKU_CURRENT_PROJECT_KEY'] = 'PBE_MASTER'
    version = get_config_version()
elif hostname == "ce-bda-dss-design-ew1-test":  # TEST DESIGN
    dataiku.set_remote_dss("http://ce-bda-dss-design-ew1-test:10000", "ytYjYLa7j9Nrh1QG6yeKkT9DJGQCzqo7")
    os.environ['DKU_CURRENT_PROJECT_KEY'] = 'PBE_MASTER'
    version = get_config_version()
elif "dss-test-standard" in hostname:  # TEST AUTOMATION
    dataiku.set_remote_dss("http://ce-bda-dss-automation-ew1-test:10000", "lzFcKW5VclB9j4QNl9UPDKxhSlsODOQP")
    os.environ['DKU_CURRENT_PROJECT_KEY'] = 'PBE_MASTER'
    version = get_config_version()
else:
    class ConfigError(Exception):
        pass
    raise ConfigError(f"Hostname '{hostname}' is not mapped to a DSS config.")
```

Additionnal settings

- Security: The API service is not public but an API key is created for webapp that should use it when calling the API. If other components need to use the API they should use nother api key.

Deployment

Actually done by data ops team.

Since API host (hostname) is important, make sure deployment follows needed naming convention (for prod, test env).

Additionnally we deploy on a kubernetes cluster with a load balancer, some settings are needed.

Steps:

1. In API Designer, Create a Version and Publish it on deployer
2. In Local Deployer, API Services, find the correct service and version, and deploy it: new deployment or update existing deployment.
3. Before the actual deployment, check the settings