

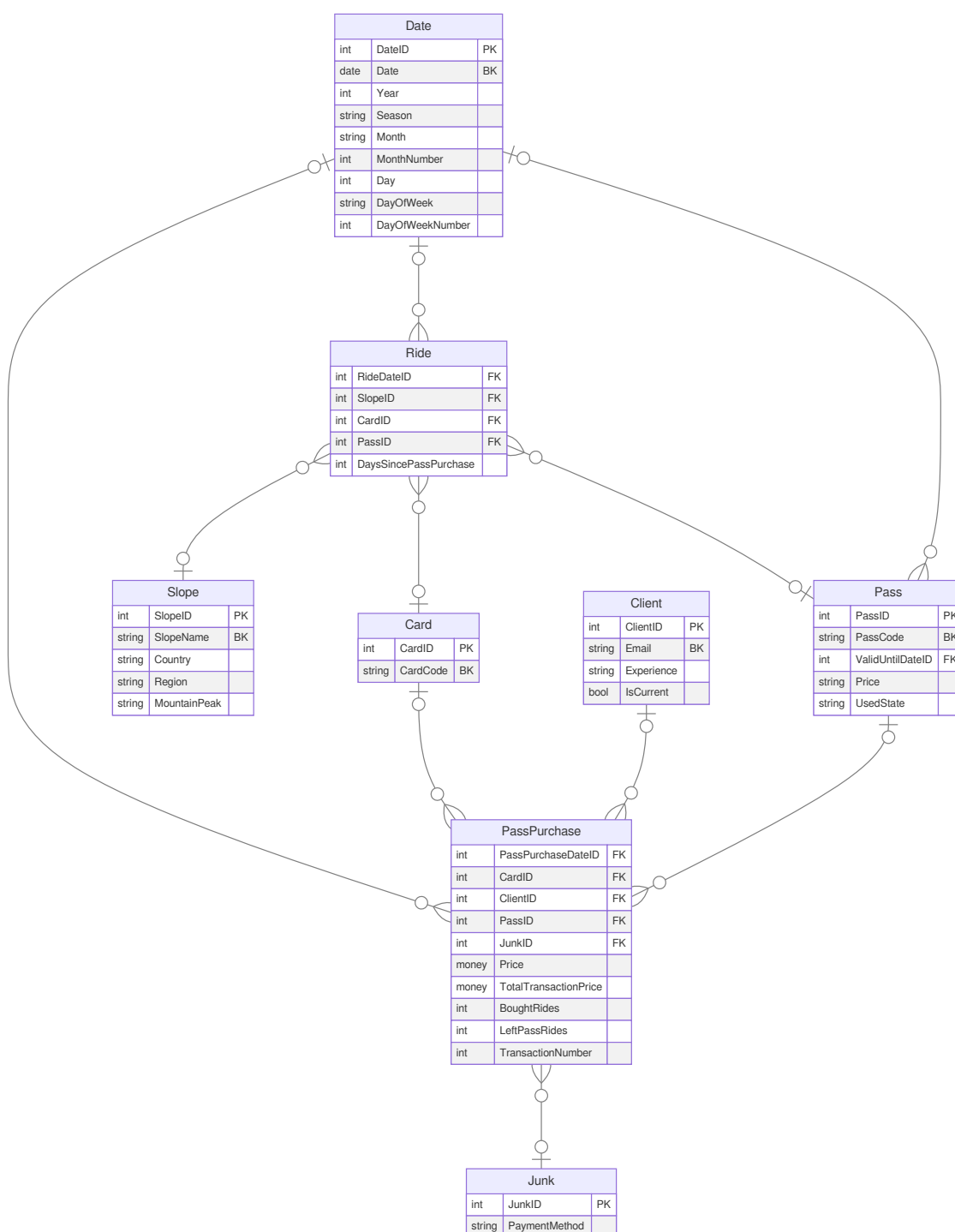
Schemat hurtowni danych dla ośrodka narciarskiego

Krzysztof Nasuta 193328, Filip Dawidowski 193433

1. Proces biznesowy

Hurtownia danych została zaprojektowana dla ośrodka narciarskiego. Opisywanym procesem biznesowym jest proces zakupu i wykorzystania kart narciarskich. Został on opisany w dokumencie pt. *Specyfikacja wymagań dla procesu biznesowego*.

2. Schemat relacyjnej bazy danych



2.1. Opis tabel

Tabela	Atrybut	Typ atrybutu	Opis
PassPurchase (Tabela faktu)	Jedna encja reprezentuje fakt wykupienia karnetu przez klienta.		
	PassPurchaseDateID	Numeric	FK Pass Data transakcji.
	CardID	Numeric	FK Card Karta, na którą został zakupiony karnet.
	ClientID	Numeric	FK Pass Klient, który zakupił karnet.
	PassID	Numeric	FK Pass Kupiony karnet.
	JunkID	Numeric	FK Junk Dodatkowe atrybuty.
	Price	Money	Cena karnetu.
	TotalTransactionPrice	Money	Całkowita kwota transakcji, której częścią jest to kupno.
	BoughtRides	Numeric	Ilość zjazdów wykupionych w karnecie.
	LeftPassRides	Numeric	Pozostała ilość zjazdów możliwych do wykonania na zakupionym karnecie.
	TransactionNumber	Numeric	Numer transakcji.
Ride (Tabela faktu)	Jedna encja reprezentuje fakt zjazdu ze stoku narciarskiego.		
	RideDateID	Numeric	FK Date Data zjazdu.
	SlopeID	Numeric	FK Slope Stok, na którym nastąpił zjazd.
	CardID	Numeric	FK Card Karta, która została wykorzystana do zjazdu.
	PassID	Numeric	FK Pass Karnet, z którego pobrano zjazd.

Tabela	Atrybut	Typ atrybutu	Opis
	DaysSincePassPurchase	Numeric	Liczba dni od zakupu karnetu do daty zjazdu.
Client (Tabela wymiaru)	Jedna encja reprezentuje pojedynczego klienta ośrodka narciarskiego.		
	ClientID	Numeric	PK (klucz zastępczy)
	Email	nvarchar(64)	Adres email klienta
	Experience	nvarchar(20)	Doświadczenie klienta na podstawie liczby wykupionych wcześniej karnetów. Dopuszczalne wartości: "1-3", "4-10", ">10"
	IsCurrent	Boolean	1 jeśli informacja jest aktualna, 0 w przeciwnym wypadku. (implementacja SCD2)
Card (Tabela wymiaru)	Jedna encja reprezentuje pojedynczą kartę zjazdową.		
	CardID	Numeric	PK (klucz zastępczy)
	CardCode	nvarchar(64)	Kod karty zjazdowej.
Date (Tabela wymiaru)	Jedna encja reprezentuje pojedynczy dzień.		
	DateID	Numeric	PK (klucz zastępczy)
	Date	Date	Data.
	Year	Numeric	Rok.
	Season	nvarchar(20)	Sezon narciarski. Dopuszczalne wartości: "Sezon XXXX", gdzie XXXX to rok rozpoczęcia sezonu.
	Month	nvarchar(20)	Nazwa miesiąca. Dopuszczalne wartości: Styczeń, Luty, Marzec, Kwiecień, Maj, Czerwiec, Lipiec, Sierpień, Wrzesień,

Tabela	Atrybut	Typ atrybutu	Opis
			Październik, Listopad, Grudzień
	MonthNumber	Numeric	Numer miesiąca. (1-12)
	Day	Numeric	Numer dnia. (1-31)
	DayOfWeek	nvarchar(20)	Dzień tygodnia. Dopuszczalne wartości: Poniedziałek, Wtorek, Środa, Czwartek, Piątek, Sobota, Niedziela
	DayOfWeekNumber	Numeric	Numer dnia tygodnia. (1-7)
Junk (Tabela wymiaru)	Encje reprezentują dodatkowe atrybuty.		
	JunkID	Numeric	PK (klucz zastępczy)
	TransactionType	nvarchar(10)	Typ transakcji. Dopuszczalne wartości: “online”, “offline”
Pass (Tabela wymiaru)	Jedna encja reprezentuje jeden karnet.		
	PassID	Numeric	PK (klucz zastępczy)
	PassCode	nvarchar(32)	Kod karnetu w formacie “Karnet XXXX”, gdzie XXXX to numer karnetu.
	ValidUntilDateID	Numeric	FK Date Data ważności karnetu.
	Price	nvarchar(20)	Cena karnetu. Dopuszczalne wartości: “0-25€”, “25-50€”, “50-100€”, “100-200€”, “200+€”
	UsedState	nvarchar(20)	Stan karnetu. Dopuszczalne wartości: “wykorzystany”, “aktywny”, “wygasły”
Slope (Tabela wymiaru)	Jedna encja reprezentuje jeden stok narciarski.		

Tabela	Atrybut	Typ atrybutu	Opis
	SlopeID	Numeric	PK (klucz zastępczy)
	SlopeName	nvarchar(20)	Nazwa stoku narciarskiego.
	Country	nvarchar(20)	Kraj stoku narciarskiego.
	Region	nvarchar(20)	Region, w którym znajduje się stok.
	MountainPeak	nvarchar(20)	Szczyt górski, na którym znajduje się stok.

3. Model wymiarowy

3.1. Definicje faktów

3.1.1. Fakt 1 - Zjazd ze stoku

Zjazd ze stoku narciarskiego, dokonany w danym dniu, na danym stoku, przy użyciu jednej karty, na której znajduje się karnet umożliwiający zjazd.

Tabela faktu: **Ride**

Ziarnistość:

- określony dzień, w którym odbył się zjazd,
- określony stok, na którym odbył się zjazd,
- określona karta, z której skorzystano,
- określony karnet, z którego pobrano zjazd.

Miary:

- Ilość zjazdów - `COUNT(*)`
- Średnia liczba zjazdów na osobę - `COUNT(*) / COUNT(DISTINCT CardID)`
- Maksymalna liczba dni od zakupu karnetu do zjazdu - `MAX(DaysSincePassPurchase)`

3.1.2. Fakt 2 - Wykupienie karnetu przez klienta

Zakup karnetu narciarskiego, dokonany w danym dniu. Zakupu dokonuje jeden klient o danym poziomie doświadczenia, w ramach jednej transakcji, dokonanej offline lub online. Kupiony karnet ma określoną cenę oraz liczbę zjazdów, do których uprawnia. Jest on przypisany do jednej karty.

Tabela faktu: **PassPurchase**

Ziarnistość:

- określony dzień transakcji,
- określony klient, który zakupił karnet,
- określona karta, do której przypisano karnet,
- określony karnet, wraz z jego ceną i ilością zjazdów,
- określona transakcja, wraz z wybranym rodzajem transakcji (offline/online).

Miary i funkcje agregujące:

- Ilość sprzedanych karnetów - $\text{COUNT}(*)$
- Średnia ilość zjazdów wykupionych w karnetach - $\text{SUM}(\text{BoughtRides}) / \text{COUNT}(*)$
- Liczba transakcji - $\text{COUNT}(\text{DISTINCT TransactionNumber})$
- Średnia liczba transakcji na osobę - $\text{Liczba transakcji} / \text{COUNT}(\text{DISTINCT ClientID})$
- Łączna kwota - $\text{SUM}(\text{Price})$
- Przychód - $\text{łączna kwota} / 1.23$
- Średnia pozostała ilość zjazdów na karnetach - $\text{SUM}(\text{LeftPassRides}) / \text{COUNT}(*)$

3.2. Definicje wymiarów

3.2.1. Wymiary dla faktu 1 - Zjazd ze stoku

Wymiar/atribut wymiaru	Tabela/column	Typ
Hierarchia daty zjazdu	<ul style="list-style-type: none"> • Date.Year • • Date.Month • • • Date.Day 	Wymiar hierarchiczny
Data zjazdu	Date	Wymiar
Rok zjazdu	Date.Year	Atrybut wymiaru
Sezon zjazdu	Date.Season	Atrybut wymiaru
Miesiąc zjazdu	Date.Month	Atrybut wymiaru
Dzień zjazdu	Date.Day	Atrybut wymiaru
Dzień tygodnia zjazdu	Date.DayOfWeek	Atrybut wymiaru
Hierarchia lokalizacji stoku	<ul style="list-style-type: none"> • Slope.Country • • Slope.Region • • • Slope.MountainPeak 	Wymiar hierarchiczny
Stok	Slope	Wymiar
Nazwa stoku	Slope.SlopeName	Atrybut wymiaru
Kraj stoku	Slope.Country	Atrybut wymiaru
Region stoku	Slope.Region	Atrybut wymiaru
Szczyt górski	Slope.MountainPeak	Atrybut wymiaru
Karta	Card	Wymiar
Kod karty	Card.CardCode	Atrybut wymiaru
Karnet	Pass	Wymiar
Kod karnetu	Pass.PassCode	Atrybut wymiaru
Cena	Pass.Price	Atrybut wymiaru
Stan karnetu	Pass.UsedState	Atrybut wymiaru

3.2.2. Wymiary dla faktu 2 - Wykupienie karnetu przez klienta

Wymiar/atribut wymiaru	Tabela/column	Typ
Numer transakcji	PassPurchase.TransactionNumber	Wymiar zdegenerowany
Hierarchia daty zakupu	<ul style="list-style-type: none"> • Date.Year • • Date.Month • • • Date.Day 	Wymiar hierarchiczny

Wymiar/atribut wymiaru	Tabela/column	Typ
Data zakupu	Date	Wymiar
Rok zakupu	Date.Year	Atrybut wymiaru
Sezon zakupu	Date.Season	Atrybut wymiaru
Miesiąc zakupu	Date.Month	Atrybut wymiaru
Dzień zakupu	Date.Day	Atrybut wymiaru
Dzień tygodnia zakupu	Date.DayOfWeek	Atrybut wymiaru
Karnet	Pass	Wymiar
Cena	Pass.Price	Atrybut wymiaru
Kod karnetu	Pass.PassCode	Atrybut wymiaru
Stan karnetu	Pass.UsedState	Atrybut wymiaru
Hierarchia daty ważności karnetu	• Date.Year • • Date.Month • • • Date.Day	Wymiar hierarchiczny
Data ważności karnetu	Date	Wymiar
Rok ważności karnetu	Date.Year	Atrybut wymiaru
Sezon ważności karnetu	Date.Season	Atrybut wymiaru
Miesiąc ważności karnetu	Date.Month	Atrybut wymiaru
Dzień ważności karnetu	Date.Day	Atrybut wymiaru
Dzień tygodnia ważności karnetu	Date.DayOfWeek	Atrybut wymiaru
Karta	Card	Wymiar
Kod karty	Card.CardCode	Atrybut wymiaru
Klient	Client	Wymiar
Email	Client.Email	Atrybut wymiaru
Doświadczenie	Client.Experience	Atrybut wymiaru
Junk	Junk	Wymiar
Typ transakcji	Junk.TransactionType	Atrybut wymiaru

4. Sprawdzenie możliwości wykonania zapytań

4.1. Oblicz średnią ilość zjazdów jednej osoby na każdym ze stoków w zależności od dnia tygodnia.

Miara	Średnia liczba zjazdów na osobę		
Wymiar	Stok	Atrybuty wymiaru	Nazwa stoku
Wymiar	Data zjazdu	Atrybuty wymiaru	Dzień tygodnia zjazdu

4.2. Ile średnio zjazdów wykonuje jedna osoba w ciągu dnia?

Miara	Średnia liczba zjazdów na osobę
-------	---------------------------------

Wymiar	Data zjazdu	Atrybuty wymiaru	Rok zjazdu, Miesiąc zjazdu, Dzień zjazdu
--------	-------------	------------------	--

4.3. Porównanie ilości sprzedanych karnetów w zależności od miesiąca.

Miara	<u>Ilość sprzedanych karnetów</u>		
Wymiar	Data zakupu	Atrybuty wymiaru	Miesiąc zakupu

4.4. Porównaj ilość karnetów zakupionych online i offline względem poprzedniego sezonu.

Miara	<u>Ilość sprzedanych karnetów</u>		
Wymiar	Data zakupu	Atrybuty wymiaru	Sezon zakupu
Wymiar	Junk	Atrybuty wymiaru	Typ transakcji

4.5. Porównaj popularność karnetów upoważniających do różnej ilości zjazdów.

Miara	<u>Ilość sprzedanych karnetów</u>		
Wymiar	Karnet	Atrybuty wymiaru	Łączna ilość zjazdów

4.6. Ile zjazdów średnio wykonuje się w ciągu miesiąca korzystając z karnetów o różnej cenie?

Miara	<u>Ilość zjazdów</u>		
Wymiar	Karnet	Atrybuty wymiaru	Cena
Wymiar	Data zjazdu	Atrybuty wymiaru	Miesiąc zjazdu

4.7. Jak długo trwa korzystanie z karnetu w zależności od jego ceny?

Miara	<u>Maksymalna liczba dni od zakupu karnetu do zjazdu</u>		
Wymiar	Karnet	Atrybuty wymiaru	Cena

4.8. Czy klienci kupujący karnety online częściej wykorzystują wszystkie zjazdy niż klienci kupujący karnety w punkcie sprzedaży?

Miara	<u>Ilość sprzedanych karnetów</u>		
Wymiar	Junk	Atrybuty wymiaru	Typ transakcji
Wymiar	Karnet	Atrybuty wymiaru	Stan karnetu

4.9. Ile średnio zjazdów pozostaje niewykorzystanych na karnetach w zależności od ich ceny?

Miara	<u>Średnia pozostała ilość zjazdów na karnecie, z którego pobrano zjazd</u>		
Wymiar	Karnet	Atrybuty wymiaru	Cena

4.10. Jak zmienia się ilość wykupionych zjazdów w zależności od doświadczenia klienta (ilości kupionych wcześniej karnetów)?

Miara	Ilość sprzedanych karnetów		
Wymiar	Karnet	Atrybuty wymiaru	Łączna ilość zjazdów
Wymiar	Klient	Atrybuty wymiaru	Doświadczenie

5. Weryfikacja dostępności wymaganych danych w źródłach

Tabela	Kolumna	Źródło danych
PassPurchase	Jedna encja reprezentuje fakt wykupienia karnetu przez klienta.	
	PassPurchaseDateID	ID daty zakupu karnetu. Klucz obcy z tabeli wymiaru. Bazuje na polu Date z tabeli Transactions z bazy danych.
	CardID	ID karty, na którą zakupiono karnet. Klucz obcy z tabeli wymiaru. Bazuje na polu CardCode z tabeli Cards z bazy danych.
	ClientID	ID klienta, który dokonał transakcji. Klucz obcy z tabeli wymiaru. Bazuje na polu Email z tabeli Clients z bazy danych.
	PassID	ID karnetu, który został zakupiony. Klucz obcy z tabeli wymiaru. Tworzony na podstawie klucza głównego PassID z tabeli Passes z bazy danych.
	JunkID	ID dodatkowych atrybutów. Klucz obcy z tabeli wymiaru. Bazuje na polu Type z tabeli Transactions z bazy danych.
	Price	Cena zakupu karnetu. Bazuje na polu Price z tabeli Passes z bazy danych.
	TotalTransactionPrice	Kwota całej transakcji, której częścią jest zakup tego karnetu. Bazuje na polu TotalPrice z tabeli Transactions z bazy danych.
	BoughtRides	Ilość zakupionych zjazdów w karnecie. Bazuje na polu TotalRides z tabeli Passes z bazy danych.
	LeftPassRides	Pozostała ilość zjazdów możliwych do wykonania na zakupionym karnecie. Bazuje na różnicy pól TotalRides i UsedRides z tabeli Passes.
	TransactionNumber	Numer transakcji. Wymiar zdegenerowany. Bazuje na polu TransactionID z tabeli Transactions z bazy danych.

Ride	Jedna encja reprezentuje fakt zjazdu ze stoku narciarskiego.	
	RideDateID	ID daty zjazdu. Klucz obcy z tabeli wymiaru. Bazuje na polu data i godzina odbicia karty z logów bramek w formacie csv.

Tabela	Kolumna	Źródło danych
	SlopeID	ID stoku, na którym odbył się zjazd. Klucz obcy z tabeli wymiaru. Bazuje na polu numer bramki z logów bramek w formacie csv.
	CardID	ID karty, którą odbito w bramce. Klucz obcy z tabeli wymiaru. Bazuje kluczu biznesowym CardCode z tabeli Cards oraz pola ID karnetu z logów bramek w formacie csv.
	PassID	ID karnetu, z którego pobrano zjazd. Klucz obcy z tabeli wymiaru. Bazuje na polu ID karnetu z logów bramek w formacie csv.
	DaysSincePassPurchase	Liczba dni od zakupu karnetu do daty zjazdu. Bazuje na różnicy pól data i godzina odbicia kart z logów bramek w formacie csv i Date z tabeli Transactions.
Client	Jedna encja reprezentuje pojedynczego klienta ośrodka narciarskiego. (Implementacja SCD2)	
	ClientID	ID klienta ośrodka narciarskiego. Klucz główny sztuczny - nadawany automatycznie.
	Email	Adres email klienta. Klucz biznesowy pobrany z pola Email z tabeli Clients z bazy danych.
	Experience	Doświadczenie klienta na podstawie liczby wykupionych wcześniej karnetów. Dopuszczalne wartości: "1-3", "4-10", ">10". Wartości mogą zostać policzone na podstawie tabel Transactions i Passes z bazy danych.
	IsCurrent	"1" jeśli informacja jest aktualna, "0" w przeciwnym wypadku. (implementacja SCD2).
Card	Jedna encja reprezentuje pojedynczą kartę zjazdową.	
	CardID	ID karty. Klucz główny sztuczny - nadawany automatycznie.
	CardCode	Kod kreskowy karty. Klucz biznesowy pobrany z pola CardCode z tabeli Cards z bazy danych.
Date	Jedna encja reprezentuje pojedynczy dzień.	
	DateID	ID daty. Klucz główny sztuczny - nadawany automatycznie.
	Wszystkie dane wymagane dla tej tabeli mogą zostać wygenerowane na podstawie dowolnego kalendarza przed procesem ETL.	
Junk	Encje reprezentują dodatkowe atrybuty.	
	JunkID	ID dodatkowych atrybutów. Klucz główny sztuczny - nadawany automatycznie.

Tabela	Kolumna	Źródło danych
Wszystkie encje odpowiadają wszystkim możliwym wartościom Type z tabeli Transactions , wygenerowane będą przed procesem ETL.		
Pass	Jedna encja reprezentuje jeden karnet.	
	PassID	ID karnetu. Klucz główny sztuczny - nadawany automatycznie.
	PassCode	Numer karnetu. Klucz biznesowy, utworzony na podstawie pola PassID z tabeli Passes .
	ValidUntilDateID	ID daty ważności karnetu. Klucz obcy z tabeli wymiaru. Bazuje na polu ValidUntil z tabeli Passes z bazy danych.
	Price	Cena karnetu. Bazuje na polu Price z tabeli Passes z bazy danych.
	UsedState	Stan wykorzystania karnetu. Dopuszczalne wartości "wykorzystany", "aktywny", "wygasły". Wartości mogą zostać policzone na podstawie tabeli Passes z bazy danych.
Slope	Jedna encja reprezentuje jeden stok narciarski.	
	SlopeID	ID stoku. Klucz główny sztuczny - nadawany automatycznie.
Dane o stokach nie ulegają zmianom, więc mogą zostać wprowadzone ręcznie przed procesem ETL.		