A 10.11 kopie
Oryginał w Archiwum Zakładowym IT Wysokie

INSTYTUT TURYSTYKI
WARSZAWA

INSTYTUT EKONOMIKI OBROTU TOWAROWEGO
AKADEMII EKONOMICZNEJ W KRAKOWIE

PRZESTRZENNE MODELE ANALIZY ROZWOJU TURYSTYKI KRAJOWEJ

(etap I)

temat III. 4.6. realizowane w ramach CPBP 08.06

„TURYSTYKA JAKO CZYNNIK ROZWOJU SPOŁECZNO-GOSPODARCZEGO"

Autorzy:

Doc. dr hab. Tadeusz Grabiński
Dr Anna Nowakowska

Kraków — Warszawa 1986

Koncepcja komputerowej

bazy danych GOSTUR

Opracowanie wykonane w ramach
CPBP 08.06 temat III.4.6 pt. "Przestrzenne
modele analizy rozwoju turystyki krajowej"

Tadeusz Grabiński
Anna Nowakowska

Kraków 1986

Spis treści

Str

1. Uwagi wstępne 3
2. Bank danych jako element systemu informatycznego 4
3. Struktura banku danych 7
4. Etapy budowy komputerowej bazy danych 11
5. Operacjonalizacja baz danych 13
6. Problem kompletności w bazach komputerowych 16
7. Metody szacowania brakujących informacji 23
8. Ogólna charakterystyka bazy danych GOSTAT 25
9. Zakres bazy danych GOSTAT 31
Słownik zmiennych pierwotnych bazy GOSTAT 33

1. Uwagi wstępne

Prowadzenie złożonych badań, w których często zachodzi konieczność korzystania z tego samego zestawu informacji w różnych etapach analizy lub też niezbędne jest powtarzanie tego samego cyklu obliczeń przy zmienionych parametrach, staje się o wiele łatwiejsze w przypadku dysponowania jednorazowo założonym, kompleksowym zbiorem danych. W sytuacji, gdy większość procedur danego systemu realizowana jest przy pomocy maszyn cyfrowych, odpowiednio skonstruowany komputerowy bank informacji w zasadniczy sposób wpływa na efektywność i dyspozycyjność całego systemu przetwarzania.

Z tego też względu przed rozpoczęciem badań w ramach tematu III.4.6 "Przestrzenne modele analizy rozwoju turystyki krajowej" wchodzącego w skład CPBP 08.06 "Turystyka jako czynnik rozwoju społeczno-gospodarczego" zdecydowano opracować komputerową bazę danych o nazwie GOSTUR, w której zawarte będą wszystkie niezbędne informacje statystyczne.

Baza ta zawierać będzie dane w układzie 49 województw Polski, dla kolejnych lat, począwszy od roku 1974, to już od momentu, w którym zaczęto publikować informacje statystyczne w nowym układzie administracyjnym. Tworzenie bazy będzie miało miejsce głównie na początku II etapu badań, tak aby dalsze prace mogły się już opierać na zebranych danych. Tym niemniej w toku 4-letnich badań przewiduje się stałą aktualizację bazy, zarówno czasową jak merytoryczną, zakładając, że docelowo cały zbiór zmiennych pierwotnych będzie liczył 300-500 pozycji.

W związku z podejmowaniem w Instytucie Ekonomiki Obrotu Towarowego Akademii Ekonomicznej w Krakowie szerokich badań

empirycznych mających na celu analizę prawidłowości zjawisk przestrzennych (w tym również w układzie województw) w sferze rynku, spożycia i konsumpcji przewiduje się sprzężenie bazy GOSTUR z bazą SPACE, która będzie tworzona równolegle dla potrzeb innych zespołów naukowo-badawczych. Pozwoli to na znaczne rozszerzenie zasobów informacyjnych stawianych do dyspozycji uczestnikom badań, przy stosunkowo niewielkich nakładach pracy.

Baza GOSTUR będzie mogła być wykorzystywana nie tylko jednorazowo do celów podjętego tematu, ale także do innych badań o charakterze przestrzennym realizowanych w przyszłości lub w innych ośrodkach.

Opracowanie niniejsze stanowi wstępną wersję pełnej koncepcji metodologicznej, która będzie podstawą konstrukcji bazy. Może ona ulec zmianie w toku realizacji prac w zależności od dostępności danych oraz technicznych środków przetwarzania. Przewiduje się dwukrotną edycję bazy - po I etapie prac oraz po ich zakończeniu (to jest po IV etapie). Druga edycja obejmować będzie również dokumentację programową oraz podręcznik korzystania z bazy.

Oprogramowanie narzędziowe bazy ukierunkowane będzie na minikomputery IBM - podobne, który to sprzęt powoli staje się w warunkach polskich standardem hardware'owym. Pozwoli to na uniknięcie kłopotów przy przekazywaniu bazy innym użytkownikom.

2. Bank danych jako element systemu informatycznego

Termin "bank danych" używany jest często zamiennie z pojęciem "baza danych". Na ogół jednak przez bank danych rozumie się tryb wykorzystywania bazy danych. Formalna definicja banku

danych mówi), że jest to wyodrębniony z systemu informatycznego podsystem realizujący funkcje tworzenia bazy danych, jej aktualizację, wybieranie i udostępnianie danych według żądanej struktury oraz ochronę danych. Bank danych stanowi bibliotekę uporządkowanych zbiorów informacji liczbowych lub tekstowych, wyposażoną w środki automatycznego wyszukiwania i [redacted] informacji oraz programy przetwarzania danych, znajdujące się w ciągłej dyspozycji użytkownika.

Do podstawowych elementów komputerowych banków danych zalicza się:

- bazy danych,
- system zarządzania bazą,
- język banku,
- programy użytkowe.

Baza danych to zbiór wzajemnie powiązanych danych przechowywany w pamięciach zewnętrznych komputera spełniający następujące warunki:

- brak redundacji (bądź jej zredukowanie do minimum) poprzez tylko jednorazowe uwzględnienie danych bądź minimalizację liczby ich powtórzeń,
- kompletność informacji w założonym zakresie przestrzennym, czasowym oraz merytorycznym,
- strukturyzacja danych wyrażająca się w łatwym i szybkim ich dostępie przez programy użytkowe oraz niskich kosztach przechowywania w pamięciach masowych,
- centralizacja zarządzania przez administratora bazy danych,
- umożliwianie wykorzystania danych równocześnie przez wielu użytkowników,
- uniezależnienie struktury danych od specyfiki programów użytkowych,

- standaryzacja rozwiązań systemowych ułatwiająca wymianę baz oraz ich konserwację (kopiowanie, aktualizację, rozszerzanie, modyfikacje),
- możliwość tworzenia zbiorów operacyjnych w dowolnych przekrojach wraz z dokonywaniem szeregu przekształceń i transformacji danych wyjściowych,
- kontrola dostępu do danych i ich ochrona przed nieupoważnionymi użytkownikami.

System zarządzania bazą jest pakietem programów organizujących obsługę użytkowników, zapewniających łączność między programami użytkowymi a systemem operacyjnym i pamięcią masową komputera oraz realizujących funkcje administrujące bazą.

W szczególności programy te związane są z:

- zakładaniem bazy danych,
- aktualizacją i rozszerzaniem zakresu danych,
- usuwaniem danych zdezaktualizowanych,
- przeorganizowaniem struktury bazy,
- wyszukiwaniem i sortowaniem żądanych informacji,
- generowaniem zbiorów operacyjnych,
- edycją wydruków,
- ochroną danych przed nieupoważnionym dostępem,
- zabezpieczeniem przed zniszczeniem bazy.

Powyższe zadania realizowane są przez 3 podstawowe zespoły funkcjonalne:

1) zarządzania danymi (tworzenie, aktualizacja i zabezpieczenie danych),
2) wyszukiwania danych (sortowanie i udostępnianie danych użytkownikom w żądanych układach),
3) moduł komunikacyjny (wykorzystanie bazy w trybie konwersacyjnym).

Język bazy jest elementem banku umożliwiającym posługiwanie się danymi zawartymi w bazie, a w szczególności kodowanie dowolnych struktur danych, deklarowanie operacji logicznych i arytmetycznych na danych pierwotnych, formułowanie przez użytkowników zadań itd.

W zależności od sposobu wykorzystywania języka w trakcie eksploatacji bazy wyróżnia się 2 koncepcje systemów zarządzania:

1) zamknięte (samodzielne), w których operuje się językiem wyspecjalizowanym, ściśle związanym ze strukturą danej bazy i nie wymagającym żadnych uzupełnień,
2) narzędziowe, w których używa się programów napisanych w powszechnie stosowanych językach programowania (FORTRAN, ALGOL, COBOL, PL) bez potrzeby wprowadzania w nich zmian synktatycznych.

Programy użytkowe w zasadzie nie stanowią elementu banku danych, lecz tworzą odrębny moduł systemu przetwarzania. Tym niemniej przydatność banku w dużej mierze uzależniona jest od ilości i rodzaju programów, które mogą bezpośrednio współpracować z bazami zawartymi w banku. W tym celu niezbędne jest takie konstruowanie "wejścia" do programów użytkowych, aby było ono kompatybilne z "wyjściem" programów kreujących zbiory operacyjne.

3. Struktura banku danych

Jak już wspomniano, bank danych składa się z kilku powiązanych modułów: bazy, programów (zarządzających oraz użytkowych) i języka - por. rys. 1. Bardzo wygodnie jest przyjąć, że bank tworzy wiele zbiorów danych o zróżnicowa-

nej strukturze zarówno z formalnego (wymiary zbioru, rodzaj informacji), jak i merytorycznego punktu widzenia (zawartość bazy). Stworzenie tylko jednego zbioru obejmującego wszystkie dane, jakkolwiek jest rozwiązaniem bardziej eleganckim, to jednak prowadzi do zbytniego skomplikowania systemu przetwarzania i związanego z tym wydłużania czasu obliczeń. Natomiast podział bazy na jednorodne podzbiory umożliwia bardziej elastyczną gospodarkę pamięcią operacyjną komputera oraz zapewnia sprawniejszą realizację celów, jakie powinien spełniać bank.

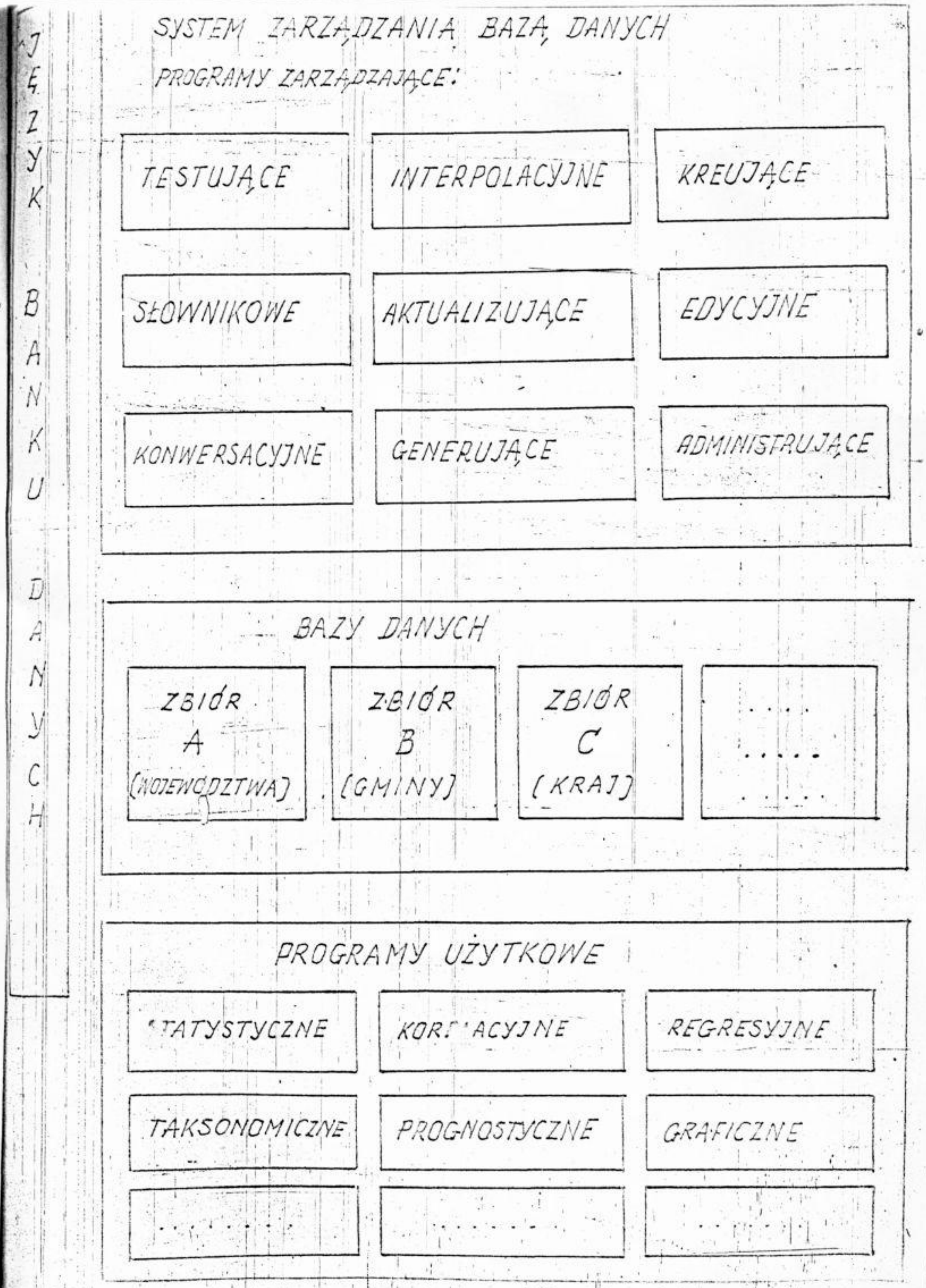
Przydatność komputerowego banku danych zależy nie tylko od ilości i jakości zawartych w nim informacji, ale także od możliwości ich przetwarzania przy pomocy odpowiednich programów obliczeniowych. Programy te wykonują różne zadania i z tego punktu widzenia można je podzielić na dwie grupy.

A) Programy zarządzające. Tworzące moduł zarządzania bazami danych i znajdujące się w gestii administratora bazy. Wyróżnić tu można przykładowo następujące rodzaje programów.

Programy testujące - przeznaczone do maszynowego sprawdzania poprawności zbiorów danych co do ich kompletności lub narzuconych ograniczeń.

Programy interpolacyjne - stosuje się w przypadku konieczności wypełnienia luk w źródłowych materiałach statystycznych. Interpolacji dokonuje się przy pomocy różnych procedur w zależności od ilości i charakteru luk, np. metody interpolacji wielomianowej, geometryczne, regresyjne, taksonomiczne itd.

Programy kreujące - transmitują dane z postaci, w jakiej znajdują się na maszynowych nośnikach informacji do postaci wymaganej przez przyjętą strukturę zbiorów w bazie.



Rys. 1. Struktura banku danych

Programy słownikowe - związane są ze specjalnymi rodzajami danych źródłowych mających charakter informacji tekstowych (wykazów, list, terminów). Są one przydatne w połączeniu z innymi programami do zwiększenia czytelności uzyskiwanych wydruków.

Programy aktualizacji - wykorzystuje się w sytuacji, gdy w trakcie eksploatacji banku uzyska się nowe materiały rozszerzające aktualny zakres danych lub też uściślające dotychczasowe informacje.

Programy edycyjne - do sporządzania wydruków zawartości bazy danych, wykazów zmiennych oraz ich wykresów w postaci kartogramów, diagramów, czy histogramów.

Programy konwersacyjne - pozwalają na pracę bezpośrednio w systemie pytanie-odpowiedź. Konstruowane są w przypadku konwersacyjnego sposobu przetwarzania danych zawartych w bazie.

Programy generujące - mają na celu tworzenie ze zbiorów bazowych dowolnych podzbiorów podlegających następnie obróbce przy pomocy programów użytkowych.

Programy administrujące - służą do manipulowania zbiorami oraz programami w ramach banku. Umożliwiają oszczędną gospodarkę pamięcią, tworzenie niezbędnych kopii zbiorów, utrzymanie porządku w banku, transmisję zbiorów między różnymi nośnikami pamięci oraz ich ochronę przed nieupoważnionym dostępem.

B) <u>Programy użytkowe</u>. Znajdujące się w gestii użytkowników banku i realizujące różnorodne zadania badawcze. Typowymi programami tego rodzaju są:

- programy statystyczne, przeznaczone do analizy badanych zbiorowości przy pomocy takich parametrów, jak: średnie, odchylenia standardowe, współczynniki zmienności, asymetrii, spłaszczenia, koncentracji, itd.;

- programy korelacyjno-regresyjne, wykorzystywane do analizy współzależności zmiennych wielowymiarowych przy użyciu współczynników korelacji liniowej i krzywoliniowej, jednorównaniowych i wielorównaniowych modeli regresji, modeli autoregresyjnych, modeli o rozłożonych opóźnieniach, ana-lizy czynnikowej, itd.;

- programy taksonomiczne, związane z analizą strukturalną mającą na celu podział zbiorowości niejednorodnych na bardziej jednorodne, dokonywanie regionalizacji, wyznaczanie wzorców rozwoju, porządkowanie obiektów według zmiennych agregatowych, dobór zmiennych diagnostycznych itd.;

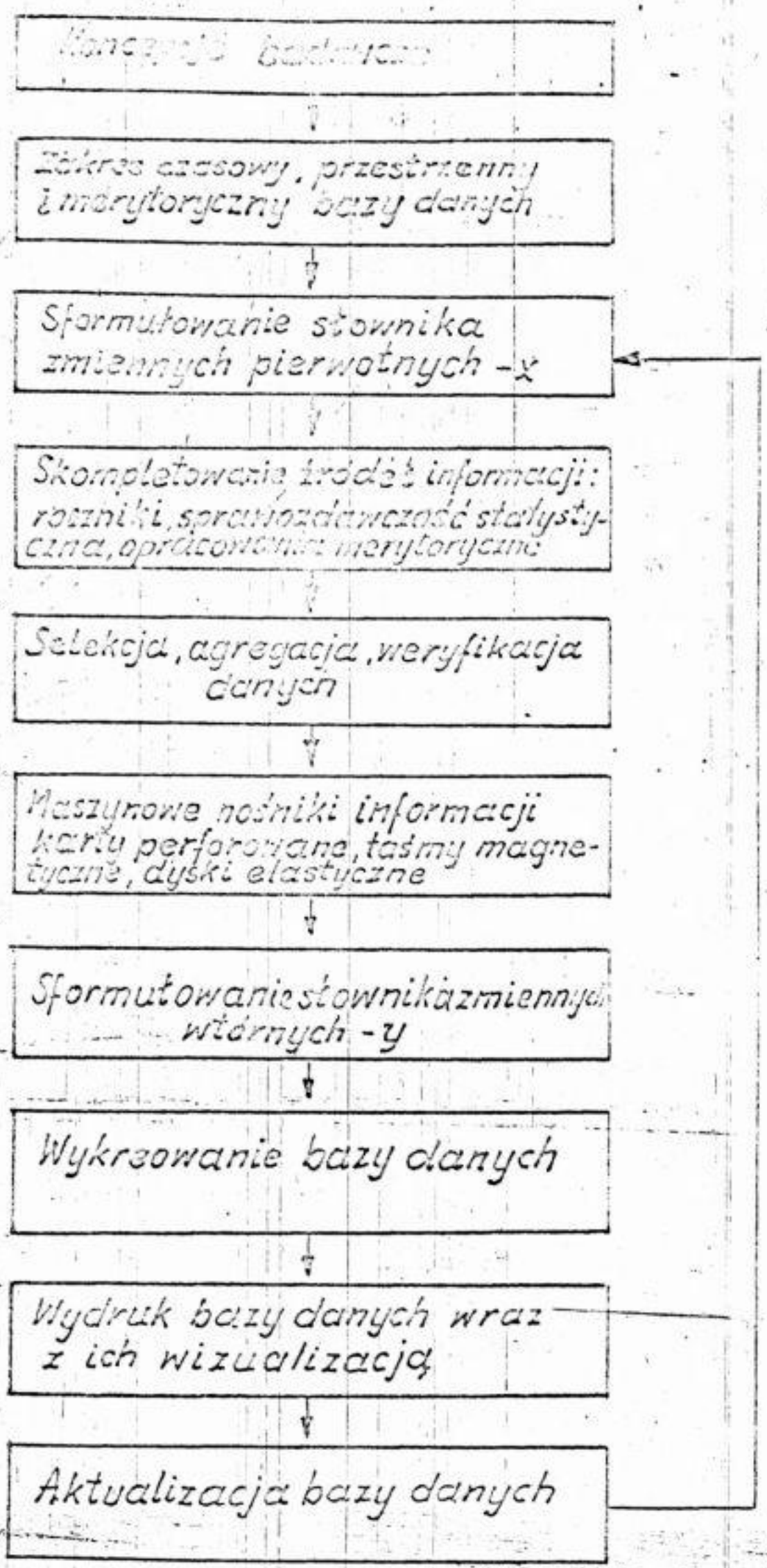
- programy prognostyczne, realizujące proces predykcji zjawisk w oparciu o takie procedury, jak ekstrapolacja funkcji trendu, analiza krzywych wzrostu, analogie rozwojowe, trendy współwiązane i prekursywne, modele regresji itd.;

- programy graficzne, stosowane w połączeniu z innymi programami lub w sposób niezależny przy sporządzaniu wykresów współzależności, wykresów tendencji rozwojowych, dendrytów, sieci powiązań, mapek konturowych, diagramów koncentracji, grawitacji i potencjału, idt.

4. Etapy budowy komputerowej bazy danych

Punktem wyjścia w procesie konstrukcji bazy danych (por. rys.2) są koncepcje badawcze determinujące charakter i konkretną postać informacji oraz ich zakres czasowy, przestrzenny i merytoryczny. Prowadzi to do sformułowania listy zmiennych pierwotnych, dla których należy zebrać informacje w ustalonym zakresie czasowym i przestrzennym.

Następny etap polega na skompletowaniu statystycznych źródeł informacji (roczników ogólnych, branżowych, sprawozdań,



Rys. 2. Etapy konstrukcji komputerowej bazy danych

opracowań merytorycznych itd.) oraz dokonaniu selekcji i agregacji zebranego materiału w celu sprawdzenia ich kompletności, rzetelności i wzajemnej porównywalności.

Jeżeli weryfikacja jakości dostępnych informacji wypadnie pozytywnie to przygotowuje się je w postaci nadającej się do dalszej obróbki komputerowej. Sprowadza się to do zakodowania danych zgodnie z instrukcjami programów kreujących lub aktualizujących oraz utworzenia maszynowych nośników informacji (kart perforowanych, dyskietek elastycznych, taśm magnetycznych).

Jeżeli natomiast jakość danych źródłowych wzbudza uzasadnione zastrzeżenia należy dokonać odpowiednich modyfikacji, wyrażających się na ogół w oszacowaniu brakujących, nierzetelnych informacji, czy też doprowadzeniu danych do porównywalności międzyobiektowej lub czasowej.

Dysponując pełną listą zmiennych pierwotnych oraz po kontroli poprawności przeniesienia ich realizacji na nośniki maszynowe należy podjąć decyzję o ewentualnym rozszerzeniu bazy o zmienne wtórne, które uzyskuje się w drodze wyznaczenia różnie zdefiniowanych wskaźników opartych na zmiennych pierwotnych. Można jednak założyć, że baza danych obejmować będzie tylko zmienne pierwotne, a tworzenie zmiennych wtórnych uwzględnić dopiero w programach generujących i pozostawić w gestii użytkowników.

Ostatni etap budowy komputerowej bazy danych polega na jej wykreowaniu, to jest na wprowadzeniu realizacji zmiennych pierwotnych bezpośrednio do pamięci komputera lub na nośniki o szybkim dostępie i większej niezawodności (dyski sztywne, taśmy magnetyczne) oraz ewentualnym dołączeniu do bazy zbioru zmiennych wtórnych. Miejsce przechowywania bazy danych zależy od jej wielkości, częstotliwości wykorzystywania przez użytkowników oraz zakresu pamięci komputera. Po wykreowaniu bazy

sporządza się pełny wydruk zawartych w niej danych wraz z ich nazwami i niekiedy wykresami ilustrującymi rozkłady wybranych zmiennych w czasie lub przestrzeni.

Budowa komputerowej bazy danych zazwyczaj nie przebiega w sposób jednorazowy i najczęściej po pewnym czasie w związku ze zmianami w koncepcji badań oraz napływem nowych informacji powstaje konieczność aktualizacji lub (i) jej modyfikacji. Proces ten odbywa się według podobnych zasad jak tworzenie nowej bazy, wymaga jednak znacznie mniejszej pracochłonności.

5. Operacjonalizacja baz danych

Typową formą organizacyjną baz danych jest kostka trójwymiarowa: obiekty x cechy x okresy:

$$X = [x_{ijt}] \qquad \begin{pmatrix} i=1,\dots,n; \\ j=1,\dots,m; \\ t=1,\dots,k \end{pmatrix} \qquad (1)$$

gdzie n, m, k, to odpowiednio liczba obiektów (jednostek przestrzennych), zmiennych oraz okresów czasu (lat).

W niektórych przypadkach poszczególne wymiary kostek danych (określone przez zbiory obiektów, cech lub okresów) mogą ulec wewnętrznemu podziałowi, tworząc macierze dwuwymiarowe typu obiekty x cechy, obiekty x okresy, cechy x okresy, obiekty x obiekty itd. lub kostki blokowe o mniejszych rozmiarach. Zwykle podziału dokonuje się na podstawie kryteriów merytorycznych i dotyczy on zbioru cech. Wyróżnione bloki danych charakteryzują w tej sytuacji różne sektory społeczno-ekonomiczne (działy lub gałęzie gospodarki narodowej, grupy tematyczne - demografia, komunikacja, handel, turystyka,

finanse itd.). Stosunkowo rzadko pojawia się natomiast potrzeba podziału zbioru obiektów (np. według kryteriów organizacyjnych, administracyjnych lub terytorialnych) względnie zbioru okresów czasu (według uprzednio dokonanej lub zadanej periodyzacji).

Istotnym problemem stojącym przed konstruktorem bazy danych jest kwestia rozszerzenia wymiarów kostki informacji. Względy ekonomiczne nakazują zbierać dane w możliwie minimalnym zakresie, podczas gdy w badaniach często zachodzi potrzeba wykorzystywania informacji pochodnych. Informacje te mogą być związane z każdym z trzech wymiarów kostki, a więc:

- okresami czasu, gdy niezbędne są informacje skumulowane za pewien okres względnie dane w postaci prognoz dla przyszłych okresów,
- obiektami, w sytuacji gdy wykorzystuje się informacje stanowiące sumę ogólną obiektów, sumę pewnych ich podzbiorów (np. dla makroregionów jako sumy województw) lub też informacje o obiektach "hipotetycznych" o przeciętnych lub wzorcowych współrzędnych),
- zmiennymi, jeżeli zachodzi potrzeba uwzględnienia w analizie dodatkowych wskaźników uzyskiwanych w drodze przekształcenia zmiennych pierwotnych zawartych w bazie.

Ogólnie biorąc możliwe są dwa podejścia w tym zakresie. W pierwszym konstruktor dostarcza użytkownikowi bazę o minimalnej redundancji, zakładając, że w trakcie eksploatacji bazy użytkownik sam będzie wyznaczał wszystkie informacje dodatkowe, jakie są mu niezbędne w badaniach. W drugim podejściu przyjmuje się, że w bazie powinny się znaleźć poza danymi pierwotnymi także i pewne, szczególnie często

spotykane w praktyce, informacje pochodne.

Zaletą pierwszego rozwiązania jest minimalne obciążenie pamięci komputera oraz duże udogodnienia na etapie kreacji bazy. Natomiast jego wadą jest konieczność wyposażenia systemu w pakiet programów umożliwiających otrzymywanie informacji pochodnych w trakcie eksploatacji bazy oraz zwiększenie czasu przetwarzania, związane z wielokrotnym obliczeniem tych samych informacji. Drugie rozwiązanie jest w znacznym stopniu pozbawione tych wad jakkolwiek nie ma ono zalet charakteryzujących rozwiązanie pierwsze.

W związku z coraz większymi możliwościami zwiększania pamięci masowych współczesnych komputerów wydaje się, że względy wygody użytkownika bazy powinny mieć priorytet. Dlatego też można przyjąć, że w komputerowych bazach poza danymi pierwotnymi wskazane jest uwzględnienie również i informacji dodatkowych, zwłaszcza związanych ze zmiennymi pochodnymi, jednakże z zachowaniem rozsądnych granic.

Przy omawianiu typowych przekształceń zmiennych możliwych do zastosowania przy generowaniu zmiennych pochodnych wygodnie jest przyjąć, iż dane wyjściowe mają charakter zmiennych naturalnych (wyrażonych w jednostkach naturalnych). Ważniejsze transformacje zmiennych naturalnych do postaci różnie zdefiniowanych wskaźników ekonomicznych przytoczono w tabl.1.

W źródłach statystycznych najczęściej spotyka się zmienne typu: A, B, C oraz F i G. Tym niemniej przy zbieraniu informacji należy zdecydować się na określoną postać zmiennych (np. zmienne naturalne typu A), natomiast wszystkie inne wskaźniki wyznaczać na ich podstawie w trakcie kreowania bazy lub w toku analiz.

Zmienne typu A wyrażone są w jednostkach naturalnych (szt., kg, zł, km^2 itd.). Zmienne typu C mają "łamane"

jednostki miary (zł/km^2, kg/osobę itd.). Pozostałe zmienne są niemianowane i po przemnożeniu ich przez 100 można je wyrazić w postaci udziałów lub wskaźników procentowych.

W badaniach społeczno-ekonomicznych zazwyczaj wykorzystuje się zmienne typu A, B, C oraz F i G. Inne typy zmiennych w praktyce spotyka się rzadko jakkolwiek mają one interesującą interpretację.

Współczynniki udziału obiektów z_{ijt} (lub okresów w_{ijt}) wskazują, jaki jest udział i-tego obiektu (t-tego okresu) w zakresie j-tej zmiennej w całym zbiorze obiektów (okresów czasu).

Współczynniki przekrojowe (czasowe) skojarzenia charakteryzują relacje między udziałami i-tego obiektu (t-tego okresu) dla zmiennych p oraz q. Jeżeli udział ten w zakresie zmiennej p jest większy niż udział w zakresie zmiennej q, to współczynniki te są większe od 1 i odwrotnie.

6. Problem kompletności w bazach komputerowych

Kompletność danych statystycznych ściśle wiąże się z problemem wiarygodności, rzetelności, dokładności i porównywalności informacji. Brak kompletności (luka) ma miejsce nie tylko w sytuacji, gdy w dostępnych źródłach statystycznych nie są dostępne odpowiednie informacje, ale również gdy istniejące dane:

- są mało wiarygodne, co można weryfikować w drodze analizy tendencji rozwojowych zmiennej, poprzez wyznaczanie wskaźników wiążących różne zmienne, czy też rozpatrując udziały poszczególnych obiektów w całym ich zbiorze,

Tablica 1

Ważniejsze przekształcenia wyjściowych zmiennych naturalnych

Typ	Postać zmiennej	definicja	Uwagi
A	wielkości naturalne	x_{ijt}	$i=1,\dots,n$ - l.obiektów $j=1,\dots,m$ - l.zmiennych $t=1,\dots,k$ - l.okresów
B	wskaźniki jednorodne	x_{ipt}/x_{iqt}	$p,q \in \{j\}$ zmienne p oraz q wyrażone są w tej samej jednostce pomiaru
C	wskaźniki niejednorodne	x_{ipt}/x_{iqt}	$p,q \in \{j\}$ zmienne p oraz q wyrażone są w różnych jednostkach pomiaru
D	współczynniki udziału obiektów z_{ijt}	$x_{ijt}/x_{.jt}$	$x_{.jt} = \sum_{i=1}^{n} x_{ijt}$
E	współczynniki udziału okresów w_{ijt}	$x_{ijt}/x_{ij.}$	$x_{ij.} = \sum_{t=1}^{k} x_{ijt}$
F	indeksy dynamiki ruchome	$x_{ij.t+1}/x_{ijt}$	$x_{ij1} = 1$
G	indeksy dynamiki stałe	x_{ijt_0}/x_{ijt}	$t_0 \in \{t\}$
H	współczynniki przekrojowe lokalizacji (skojarzenia)	z_{ipt}/z_{iqt}	$p, q \in \{j\}$ z_{ijt} - współczynnik i-tego obiektu
I	współczynniki czasowe lokalizacji (skojarzenia)	w_{ipt}/w_{iqt}	$p, q \in \{j\}$ w_{ijt} - współczynnik udziału t-tego okresu

- różnią się między sobą w zależności od źródła statystycznego,
- nie są porównywalne wskutek różnych przyczyn np. zmiany podstawy cen, przynależności administracyjnej obiektów, zmiany jednostek pomiarowych, czy też zakresu pojęciowego zmiennej,
- pochodzą ze źródeł statystycznych zarówno oficjalnych jak i nieoficjalnych.

W wymienionych powyżej sytuacjach istnieje wprawdzie możliwość wprowadzenia do bazy odpowiednich informacji jednakże wymaga to dokonywania oceny stopnia ich wiarygodności, doprowadzenia do możliwie największej porównywalności, czy też wyboru informacji najbardziej rzetelnej. W rezultacie sytuacje te niewiele różnią się od przypadku zupełnego braku informacji, kiedy w drodze analizy statystyczno-ekonomicznej można oszacować brakujące dane zapewniając w ten sposób ich kompletność.

Rozpatrując problem szerzej należy rpwnież zwrócić uwagę na fakt, że także i oficjalnie publikowane dane charakteryzują się różnym stopniem dokładności. Źródła błędów, zniekształceń, a nawet świadomych zafałszowań informacji statystycznych są rozliczne i trudno jest określić rząd ich wpływu na jakość tzw. danych oficjalnych. Spośród wielu możliwych przyczyn złej jakości danych można wymienić np.:

- nieprecyzyjne określenie jednostek badania lub jednostek sprawozdawczych,
- braki w zakresie metodologii definiowania ewidencjonowanych zjawisk,
- niespójne klasyfikacje i zasady grupowania jednostek,
- błędy nieświadome lub świadome w trakcie sporządzania pierwotnych sprawozdań statystycznych,
- zniekształcenia w czasie przenoszenia danych ze sprawozdań na maszynowe nośniki informacji, a niekiedy nawet w czasie

przetwarzania na EMC,

- błędy wynikające z braku dokładnej korekty publikacji statystycznych.

Reasumując można stwierdzić, że podejmując badania empiryczne oparte na stosunkowo dużym materiale statystycznym należy się liczyć z faktem zróżnicowanej rzetelności danych. Może to wyrażać się w obiektywnym stopniu niedokładności danych oficjalnych, konieczności weryfikacji informacji niewiarygodnych, sprzecznych lub nieporównywalnych, względnie dokonywaniu subiektywnych, własnych szacunków brakujących danych.

Dlatego też konstruktor komputerowej bazy danych statystycznych powinien na wstępie określić swoje podejście do problemu kompletności danych. Decyzje w tym zakresie zależą od rodzaju danych (np. statystyki międzynarodowe charakteryzują się stosunkowo dużą liczbą luk, w odróżnieniu od statystyk krajowych), od przewidywanego sposobu eksploatacji bazy (baza może spełniać tylko funkcje informatora względnie stanowić bezpośrednie źródło danych do programów obliczeniowych), czy wreszcie od możliwości czasowych i finansowych administratora bazy (utworzenie kompletnej bazy danych wymaga większych środków niż mechaniczne przeniesienie danych ze źródeł pisanych do pamięci komputera w odpowiednio zorganizowanej formie).

Możliwe sposoby rozwiązania problemu kompletności danych można jak się wydaje sprowadzić do trzech sytuacji.

W pierwszej przyjmuje się, że administrator bazy nie powinien dokonywać żadnych zmian ani też uzupełnień w oficjalnych danych statystycznych. W rezultacie prowadzi to na ogół do bazy danych z lukami. Przyszły użytkownik takiej bazy, wykorzystujący ją do badań statystycznych ma do wyboru 3 możliwości:

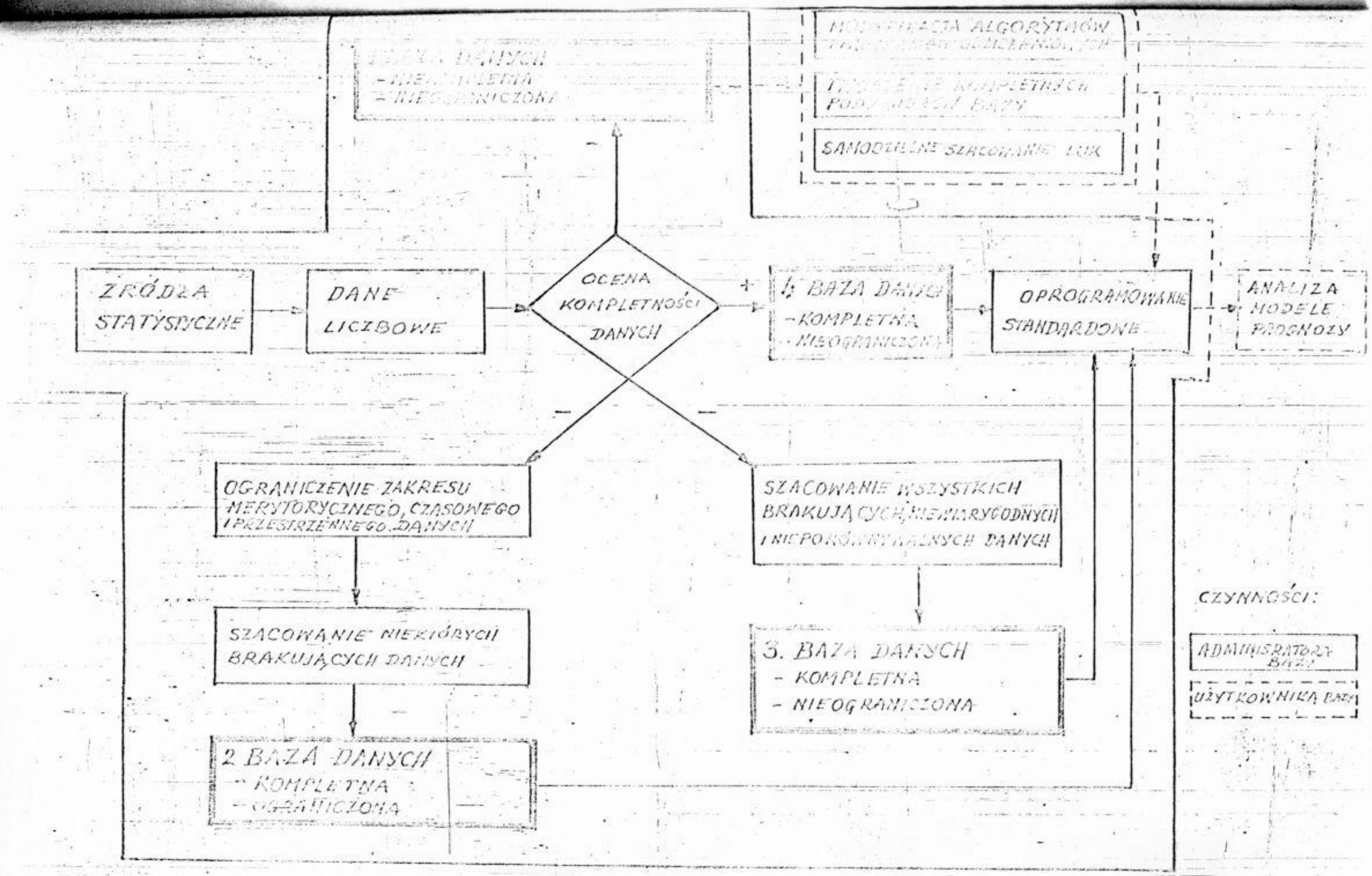
- stosować tylko te metody analizy, które zezwalają na operowanie zbiorami niekompletnymi, co jednak wymaga opracowania nowych względnie modyfikacji istniejących algorytmów i programów obliczeniowych,
- "wycinać" z bazy pewne jej fragmenty, w ramach których istnieją wszystkie informacje,
- dokonywać na własną odpowiedzialność niezbędnych szacunków brakujących danych.

W pozostałych dwóch podejściach przyjmuje się, że celem administratora jest dostarczanie kompletnej bazy danych. Rozwiązanie pośrednie zakłada, że już w etapie projektowania bazy eliminuje się te obiekty, zmienne lub okresy, dla których liczba luk w źródłach statystycznych jest stosunkowo duża, względnie dostępne informacje budzą duże zastrzeżenia co do ich wiarygodności i porównywalności. Dopuszcza się tu możliwość dokonywania szacunków oraz modyfikacji danych, jednakże w ograniczonej skali.

Trzecie krańcowe rozwiązanie polega na zniesieniu ograniczeń co do skali dokonywanych szacunków i prowadzi do uzyskiwania pełnych baz danych statystycznych bez zmniejszenia ich zakresu merytorycznego, czasowego, czy przestrzennego.

W ostatnich dwóch przypadkach niezbędne jest wyposażenie bazy w komentarz informujący użytkownika o liczbie, lokalizacji i charakterze dokonanych szacunków, jak również o wykorzystanej w tym celu metodzie badawczej. Istotnym problemem jest też określenie granicy stopnia kompletności i wiarygodności zmiennych, poniżej której ich włączenie do bazy mija się z celem z uwagi na możliwy do uzyskania rząd dokładności szacunków.

Powyższe rozważania ilustruje rys.3, gdzie obok wyróżnionych 3 sytuacji podano również czwartą, w której nie zachodzi



5.3. Kompletność danych w procesie budowy komputerowych baz danych

potrzeba uzupełniania brakujących danych ani ograniczania zakresu bazy. Zbiorczą charakterystykę omówionych sposobów z punktu widzenia trzech kryteriów: 1) stopnia kompletności danych, 2) wielkości niezbędnych szacunków oraz 3) zakresu wynikowej bazy przedstawia tablica 2.

Tablica 2

Charakterystyka możliwych sposobów rozwiązania problemu kompletności w komputerowych bazach danych statystycznych

Udział szacunków / Kompletność danych	Zakres bazy danych		
	Pełny		ograniczony
	brak	duży	częściowy
zupełna	4	3	2
częściowa	1	-	-

Każde z przedstawionych rozwiązań (z wyjątkiem czwartego, które w praktyce występuje rzadko) ma swoje wady i zalety. Rozwiązanie pierwsze jest najłatwiejsze dla konstruktora bazy, powoduje jednak u użytkownika zwiększenie pracochłonności na etapie oprogramowania algorytmów i procedur, a niekiedy wymaga zrezygnowania z niektórych metod analizy. Pozostałe dwa rozwiązania dostarczają baz danych wygodnych w eksploatacji, kosztem dużego nakładu pracy przy ich budowie. Rozwiązanie drugie jest mniej pracochłonne, ale prowadzi do baz danych o ograniczonym zakresie. Rozwiązanie trzecie jest najbardziej pracochłonne, dostarcza kompletną bazę danych w pełnym zakresie, jakkolwiek zawierającą stosunkowo dużą ilość danych szacunkowych.

7. Metody szacowania brakujących informacji

Jak łatwo zauważyć w procesie konstrukcji komputerowych baz danych statystycznych istotnym problemem jest kwestia możliwości szacowania brakujących informacji. Niestety zagadnienie to w statystyce jest rzadko podejmowane. W szczególności brak jest wyczerpujących badań nad przydatnością różnych metod interpolacyjnych.

W literaturze spotkać można szereg procedur wykorzystywanych do szacowania brakujących danych, zwłaszcza w opracowaniach z zakresu prognozowania statystycznego. Przez prognozę można bowiem rozumieć nie tylko sąd o wielkości zjawiska w przyszłości, ale także ogólniej, o nieznanej wielkości zjawiska bez względu na okres prognozy.

Przegląd procedur interpolacyjnych prowadzi do wniosku, że w większości przypadków dotyczą one szeregów czasowych. Można tu wymienić między innymi następujące metody:

1. Metody interpolacji wielowymiarowej oparte na wzorach Lagrange'a, Newtona, Gaussa, Stirlinga, Bessela, czy Aitkena, przy czym z uwagi na zazwyczaj nierównomierne rozłożenie znanych wartości szeregu na osi zmiennej czasowej (węzłów interpolacji) najbardziej przydatnym jest tu wzór interpolacyjny Lagrange'a.
2. Metody oparte na klasycznych funkcjach trendu (o różnych postaciach analitycznych), których parametry szacowane są metodą najmniejszych kwadratów w oparciu o znane wartości szeregu czasowego.
3. Metody geometryczne sprowadzające się do łączenia sąsiadujących ze sobą [redacted] dwóch lub trzech znanych elementów szeregu

czasowego, za pomocą odcinków linii prostej (przy 2 elementach) lub łuków paraboli (dla 3 elementów), a następnie określeniu według wzorów z geometrii analitycznej współrzędnych nieznanych elementów szeregu czasowego, zawartych wewnątrz danego odcinka lub łuku.

Bardziej skomplikowany problem stanowi szacowanie luk w szeregach przekrojowych. Wynika to głównie z trudności związanych z określeniem argumentu funkcji interpolacyjnej. W przypadku szeregów czasowych rolę argumentu odgrywa zmienna czasowa o znanych wartościach. Dla szeregów przekrojowych należy w tym celu korzystać z informacji zawartych w innych cechach, dla których są znane wszystkie realizacje. Cechy te, nazywane dalej wiodącymi - poza kompletnością - powinny charakteryzować się wysokim skorelowaniem ze zmienną interpolowaną.

Wyniki szacunków brakujących danych w szeregach przekrojowych różnią się nie tylko w zależności od wykorzystanej metody interpolacji, ale także od przyjętej cechy wiodącej oraz sposobu ustalenia na jej podstawie argumentu funkcji interpolacyjnej. Za argumenty te można mianowicie przyjmować:

a) numery kolejne zmiennej interpolowanej wynikające z uporządkowanych malejąco lub rosnąco wartości zmiennej wiodącej,
b) wielkości proporcjonalne do różnic między sąsiednimi wartościami uporządkowanej zmiennej wiodącej, przyjmując za normę np. maksymalną lub minimalną różnicę,
c) wartości samej zmiennej wiodącej.

W przypadku dysponowania kilkoma zmiennymi wiodącymi istnieje też możliwość dokonywania szacunków na podstawie realizacji zmiennej syntetycznej, określonej na podstawie wyjściowych zmiennych wiodących zgodnie z regułami wielowymiarowej analizy

porównawczej.

Jeśli chodzi o metody interpolacji danych w szeregach przekrojowych to można stosować tu identyczne procedury obliczeniowe, jak przy szeregach czasowych, uwzględniając jeden z wymienionych wyżej sposobów określenia argumentów funkcji interpolacyjnej. Trzeba zauważyć, że metoda klasycznych funkcji trendu (2) z przyjęciem jako argumentu tej funkcji - wartości zmiennej wiodącej (c) sprowadza się do wykorzystania funkcji regresji do celów prognozy brakujących danych.

Specyficzną dla szeregów przekrojowych metodą szacowania luk jest metoda bikwantylowa, szczególnie przydatna w przypadku dysponowania wielowymiarowym zbiorem zmiennych wiodących. Polega ona na podziale elementów szeregu przekrojowego (obiektów) na jednorodne grupy w oparciu o realizacje zmiennych wiodących, a następnie oszacowaniu luk w zmiennej interpolowanej na poziomie np. średniej arytmetycznej lub mediany ze znanych realizacji tej zmiennej dla obiektów znajdujących się w tej samej grupie.

Warto też badmienić, że na ogół dostępne informacje w bazie mają charakter szeregów przekrojowo-czasowych, w związku z czym można konstruować procedury interpolacyjne, łączące w sobie podejścia właściwe dla obu rodzajów szeregów.

3. Ogólna charakterystyka bazy danych GOSTUR

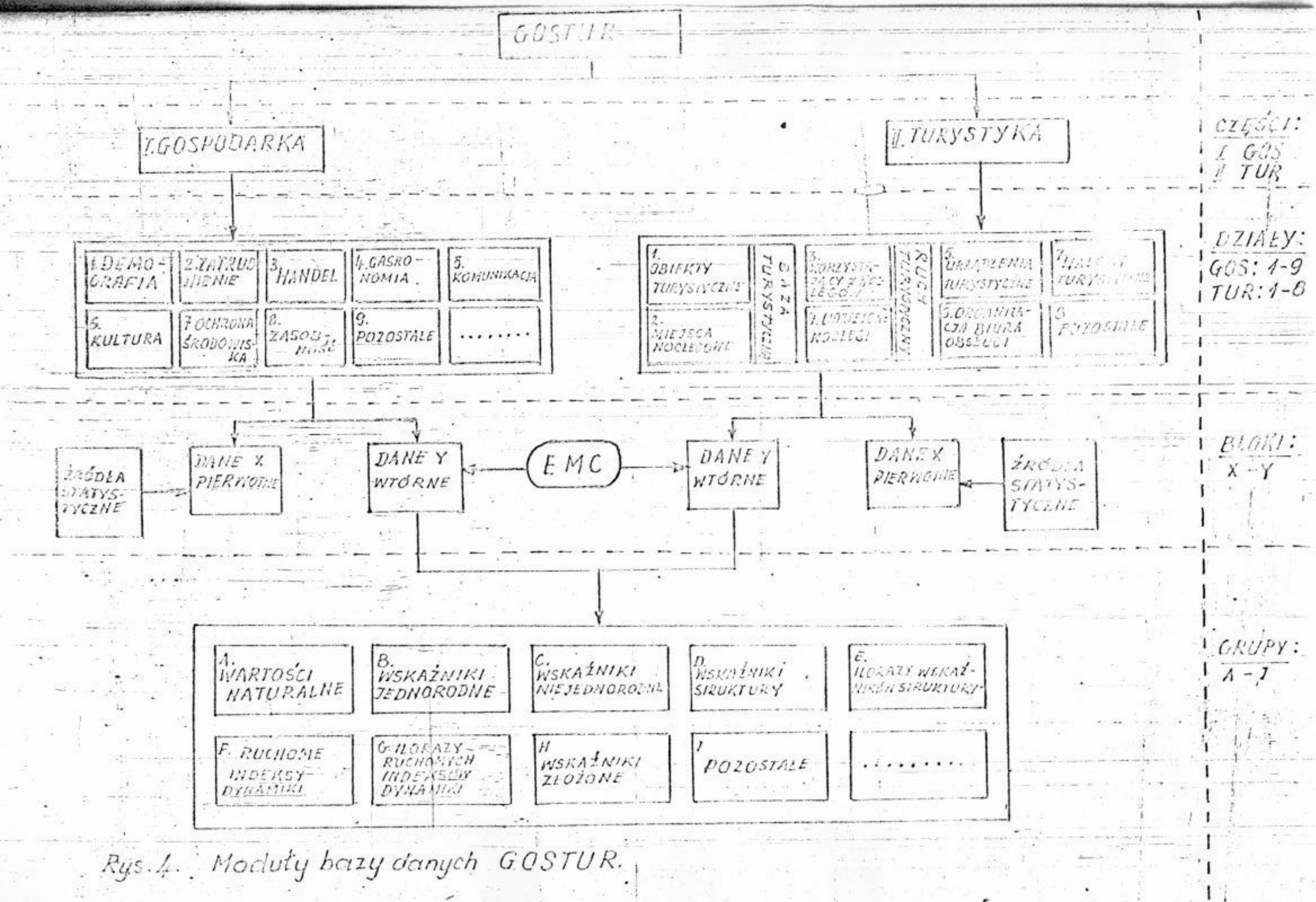
Zgodnie z ogólną definicją, komputerową bazę danych GOSTUR określić można jako element banku danych przechowywany w pamięciach zewnętrznych komputera i zawierający logicznie powiązany zbiór danych statystycznych zorganizowany w postaci kostki trójwymiarowej: zmienne x lata x województwa. Baza ta ma charakter ukierunkowany o ograniczonym zasięgu. Tworzona jest na podsta-

wie oficjalnych źródeł statystycznych (Roczniki Województw oraz Opracowania Statystyczne GUS) i wykorzystywana będzie wielokrotnie w dłuższym okresie czasu w celu przeprowadzenia wszechstronnej analizy prawidłowości rozwojowych turystyki krajowej w ujęciu przestrzennym, w powiązaniu z wybranymi elementami gospodarki narodowej.

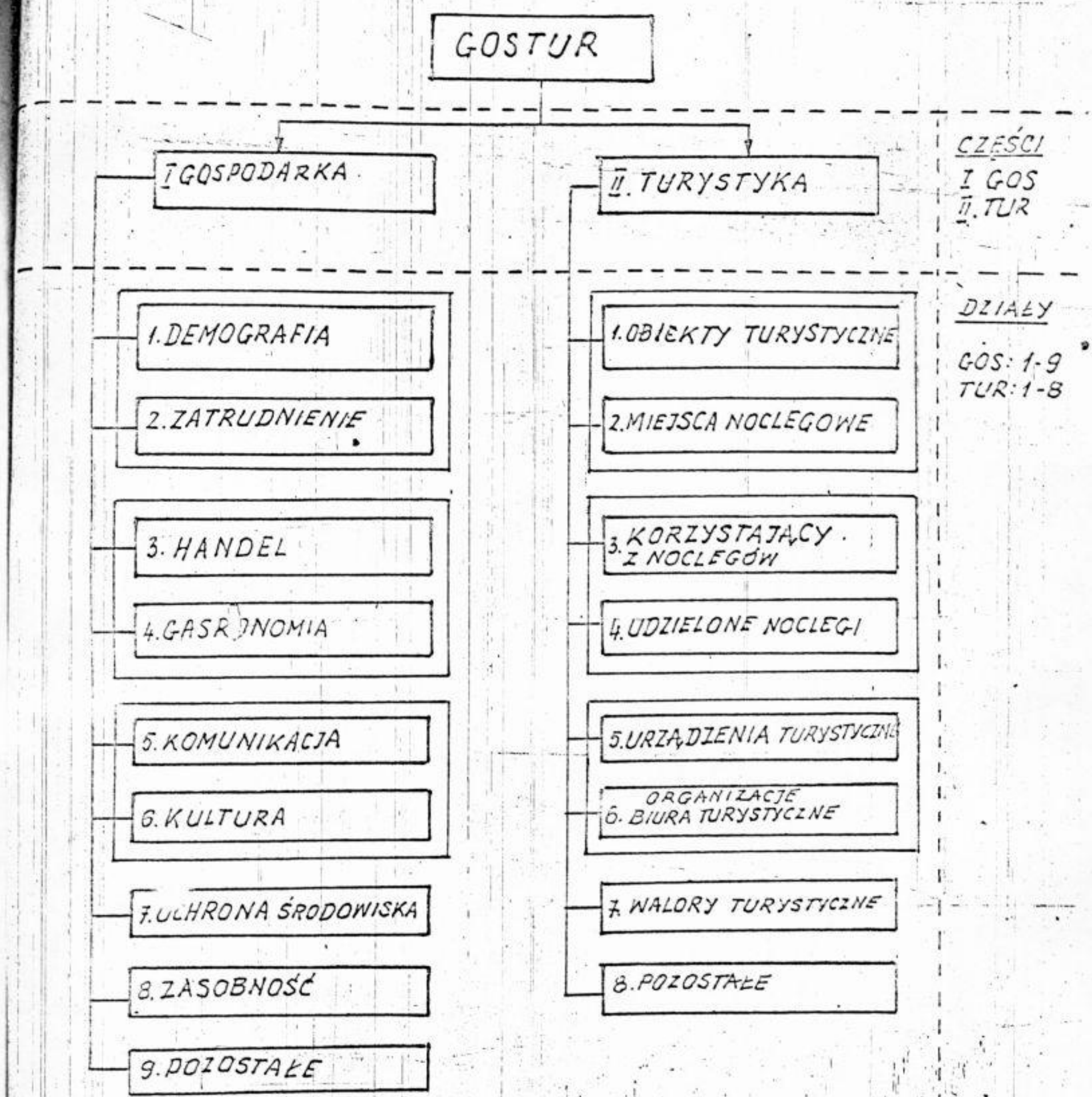
Baza GOSTUR zbudowana jest w sposób modułowy, przy czym można w niej wyróżnić 4 zasadnicze przekroje klasyfikacyjne danych (por. rys.4):

1) <u>części</u>: gospodarczą (GOS) oraz turystyczną (TUR),
2) <u>działy</u>: tematyczne w ramach każdej części (wstępnie 9 działów w części gospodarczej oraz 3 działów w części turystycznej),
3) <u>bloki</u>: zmiennych pierwotnych X oraz zmiennych wtórnych (pochodnych, wynikowych) Y,
4) <u>grupy</u> zmiennych o podobnej definicji (A-I) zgodnie z wykazem w tabl.1, przy czym podział na grupy dokonywany jest tylko w bloku zmiennych wtórnych.

Aktualny wykaz działów tematycznych wraz z ich numeracją w ramach obydwóch części bazy podany jest na rys.5. Kwalifikacja zmiennych do poszczególnych działów odbywa się na podstawie ich znaczenia merytorycznego. W miarę rozbudowy bazy wzrastać będzie również liczba jej działów tematycznych.
W każdej części bazy GOSTUR przewidziany jest odrębny dział zmiennych pomocniczych (pozostałych), w których mogą występować identyczne zmienne, co stanowi pewne odstępstwo od postulatu braku redundacji. Dzięki temu każdą z części bazy danych można traktować jako samodzielną bazę o mniejszych rozmiarach. Pozwala to znacznie zmniejszyć wielkość niezbędnej pamięci operacyjnej EMC, jeżeli tylko zakres badań nie wymaga wykorzystywania zmiennych z obu części bazy.



Rys. 4. Moduły bazy danych GOSTUR.



Rys. 5. Części i działy tematyczne bazy danych GOSTUR

Aktualny wykaz grup zmiennych wraz ze symbolami służącymi do ich identyfikacji przytoczony jest na rys.6. Podane tu grupy zmiennych nie wyczerpują wszystkich możliwości w zakresie konstruowania wszystkich współczynników, a jedynie tylko te, które są najczęściej wykorzystywane w badaniach i z tego względu zostały włączone do bazy. Inne współczynniki można wyznaczać w trakcie generowania zbiorów operacyjnych.

Podobnie jak w przypadku części bazy (GCS oraz TUR) tak i każdy blok (X oraz Y) oraz grupa zmiennych (A-I) może stanowić samodzielną bazę o mniejszych rozmiarach. Jedynie działy tematyczne zawsze ujmowane są łącznie. W zależności od potrzeb użytkownika w obliczeniach wykorzystywany jest tylko odpowiedni moduł bazy.

Dla administratora bazy danych najważniejszą są bloki zmiennych pierwotnych $\{X\}$, które wymagają ciągłej aktualizacji, doprowadzenia do wzajemnej porównywalności oraz dokonywania interpretacji brakujących informacji. Bloki zmiennych wtórnych $\{Y\}$, na które składają się poszczególne grupy zmiennych $\{A\}$ - $\{I\}$ kreowane są przy pomocy programów obliczeniowych w momencie zakładania bazy. Zadaniem konstruktora bazy jest w tym przypadku opracowanie definicji wszystkich zmiennych wtórnych, tak aby zapewnić użytkownikowi maksimum wygody w trakcie eksploatacji bazy.

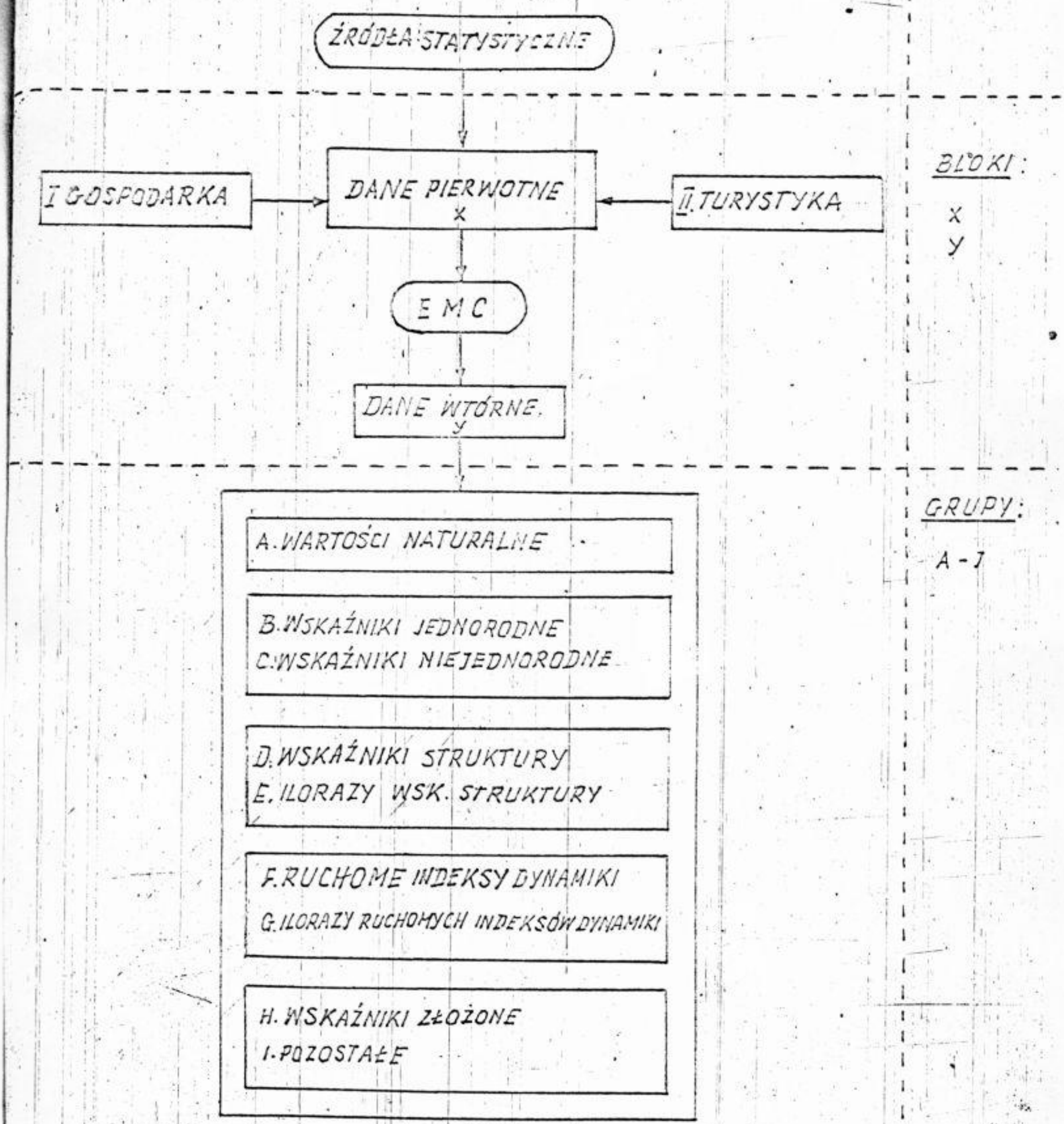
Typową formą zapisu danych na pamięciach masowych jest multifile na taśmie magnetycznej, którego poszczególnymi składowymi są kolejne moduły elementarne:

$$\{Y\} = \{A\} \cup \{B\} \cup \ldots \{I\}$$

wraz z modułem $\{X\}$ dla każdego bloku $\{GCS\}$ i $\{TUR\}$.

Moduły te można ujmować w postaci:

- pojedynczych rekordów o m x n x k elementów,



Rys. 6. Bloki i grupy zmiennych bazy danych GOSTUR

- m rekordów o n x k elementach,
- m x n rekordów o k elementach,

gdzie m, n, k to odpowiednio liczba zmiennych w danym module, liczba województw (49) oraz liczba lat. Parametry n oraz k są jednakowe dla wszystkich modułów, natomiast parametry m przyjmują dla każdego modułu inne wartości. W przypadku łączenia modułów zmiennie ulega tylko parametr m, który jest sumą parametrów m dla modułów składowych.

Wybór sposobu zapisu danych zależy od konfiguracji systemu, w którym instalowana jest baza danych. Sekwencyjna forma zapisu jest stosowana tylko do sporządzania źródłowej wersji bazy dla celów archiwizacji, aktualizacji, kopiowania itd. W przypadku przetwarzania danych wykorzystuje się zapis bazy na pamięciach dyskowych o bezpośrednim dostępie, dzięki czemu czas obliczeń jest znacznie krótszy.

9. Zakres bazy danych GOSTUR

Aktualnie tworzona baza zawiera informacje w przekroju 49 województw dla 12 lat od 1974 do 1985 roku. Zgodnie z załączonymi słownikami zmiennych pierwotnych wersja początkowa bazy obejmuje 77 zmiennych w części gospodarczej oraz 24 zmienne w części turystycznej. Tak więc łącznie w bazie znajdzie się 60 tys. informacji liczbowych.

Uwzględniając zmienne wtórne zarówno w postaci naturalnej (np. dane dla wsi jako różnica między danymi ogólnymi a danymi dla miast) jak i w postaci wskaźników (na 1 mieszkańca, na 1 km^2, udziały procentowe województw, wskaźniki złożone) pozwala to utworzyć bazę zawierającą 5-6 krotnie większą liczbę zmiennych (ok.600) to jest

ponad 300 tys. danych liczbowych.

Jak łatwo zauważyć w obecnej wersji bazy nie ma wielu zmiennych, które mogą znaleźć zastosowanie w planowanych badaniach. W szczególności w ogóle brak jest zmiennych z działów 5-7 w części turystycznej. Rozszerzenie bazy nastąpi w II etapie prac, w ramach którego sporządzony zostanie pełny wydruk zawartości bazy wraz z kompletnym słownikiem zmiennych pierwotnych oraz zmiennych wtórnych.

Słownik zmiennych pierwotnych bazy GOSTUR

(Część I. GOSpodarka)

1. Demografia

G1.1 ludność ogółem (tys.)

G1.2 ludność w wieku produkcyjnym (tys.)

G1.3 ludność z wykształceniem wyższym (tys.)

G1.4 ludność w miastach (tys.)

G1.5 ludność w miastach pow. 50 tys. mieszkańców (tys.)

2. Zatrudnienie

G2.1 zatrudnienie w gosp. uspołecznionej (tys.)

G2.2 zatrudnienie w przemyśle (tys.)

G2.3 zatrudnienie w budownictwie (tys.)

G2.4 zatrudnienie w handlu (tys.)

G2.5 zatrudnienie w gospod. uspołecznionej z wykształceniem wyższym (tys.)

3. Handel

G3.1 punkty sprzedaży detalicznej w handlu rynkowym

G3.2 punkty sprzedaży detalicznej w handlu uspołecznionym

G3.3 punkty sprzedaży detalicznej w handlu uspołecznionym w miastach

G3.4 sklepy w handlu rynkowym

G3.5 sklepy w handlu uspołecznionym

G3.6 sklepy w handlu uspołecznionym w miastach

G3.7 powierzchnia sklepów w handlu rynkowym (tys. m^2)

G3.8 powierzchnia sklepów w handlu uspołecznionym (tys. m^2)

G3.9 powierzchnia sklepów w handlu uspołecznionym w miastach (tys. m^2)

G3.10 sprzedaż detaliczna w handlu rynkowym (mln zł)

G3.11 sprzedaż detaliczna w handlu uspołecznionym (mln zł)

G3.12 sprzedaż detaliczna w handlu uspołecznionym w miastach (mln zł)

4. Gastronomia

G4.1 zakłady gastronomiczne ogółem
G4.2 zakłady gastronomiczne uspołecznione
G4.3 zakłady gastronomiczne uspołecznione w miastach
G4.4 zakłady gastronomiczne uspołecznione w miastach sezonowe
G4.5 miejsca konsumenckie w uspołecznionych zakładach gastronomicznych (tys.)
G4.6 miejsca konsumenckie w uspołecznionych zakładach gastronomicznych w miastach (tys.)
G4.7 miejsca konsumenckie w uspołecznionych zakładach gastronomicznych w miastach sezonowe (tys.)
G4.8 sprzedaż w uspołecznionych zakładach gastronomicznych (mln zł)
G4.9 sprzedaż w uspołecznionych zakładach gastronomicznych w miastach (mln zł)

5. Komunikacja

G5.1 linie kolejowe eksploatowane (km)
G5.2 drogi publiczne o twardej nawierzchni (km)
G5.3 drogi publiczne pokryte siecią PKS (km)
G5.4 liczba przystanków PKS
G5.5 liczba dworców autobusowych
G5.6 liczba placówek pocztowo-telekomunikacyjnych
G5.7 liczba placówek pocztowo-telekomunikacyjnych w miastach
G5.8 abonenci telefoniczni (tys.)
G5.9 abonenci telefoniczni w miastach (tys.)

6. Kultura

G6.1 liczba teatrów i instytucji muzycznych
G6.2 liczba teatrów
G6.3 liczba miejsc na widowni
G6.4 liczba miejsc w teatrach
G6.5 liczba przedstawień i koncertów
G6.6 liczba przedstawień w teatrach
G6.7 liczba widzów i słuchaczy (tys.)
G6.8 liczba widzów w teatrach (tys.)
G6.9 liczba muzeów
G6.10 liczba zwiedzających muzea (tys.)
G6.11 liczba kin
G6.12 liczba miejsc na widowniach kin

7. Ochrona środowiska

G7.1 liczba zakładów uciążliwych
G7.2 emisja zanieczyszczeń pyłowych (tys.t)
G7.3 emisja zanieczyszczeń gazowych (tys.t)
G7.4 łącznie zanieczyszczenia pyłowe (tys.t)
G7.5 łącznie zanieczyszczenia gazowe (tys.t)
G7.6 odpady uciążliwe nagromadzone (tys.t)
G7.7 odpady uciążliwe wytworzone (tys.t)
G7.8 odpady uciążliwe wykorzystane gospodarczo (tys.t)
G7.9 powierzchnia parków narodowych (ha)
G7.10 powierzchnia lasów ochronnych (tys.ha)
G7.11 liczba rezerwatów przyrody
G7.12 powierzchnia rezerwatów przyrody (ha)
G7.13 liczba pomników przyrody

8. Zamożność

G8.1 liczba samochodów osobowych prywatnych (tys.)
G8.2 liczba samochodów ciężarowych prywatnych (tys.)
G8.3 liczba motocykli i skuterów prywatnych (tys.)
G8.4 liczba abonentów radiowych (tys.)
G8.5 liczba abonentów radiowych w miastach (tys.)
G8.6 liczba abonentów TV (tys.)
G8.7 liczba abonentów TV w miastach (tys.)
G8.8 wkłady oszczędnościowe w PKO (mln zł)
G8.9 składki z tytułu ubezpieczeń majątkowych i osobowych (mln zł)

9. Pozostałe

G9.1 powierzchnia (km^2)
G9.2 liczba miast
G9.3 liczba gmin

Słownik zmiennych pierwotnych bazy COSTUR
(część II. TURystyka)

1. Obiekty

T1.1 liczba obiektów turystycznych ogółem

T1.2 liczba obiektów turystycznych całorocznych

T1.3 liczba obiektów wczasowo-wypoczynkowych ogółem

T1.4 liczba obiektów wczasowo-wypoczynkowych całorocznych

2. Miejsca noclegowe

T2.1 liczba miejsc w obiektach turystycznych ogółem

T2.2 liczba miejsc w obiektach turystycznych całorocznych

T2.3 liczba miejsc w obiektach turystycznych w dniu 31.7 ogółem

T2.4 liczba miejsc w obiektach turystycznych w dniu 31.7 całorocznych

T2.5 liczba miejsc w obiektach wczasowo-wypoczynkowych ogółem

T2.6 liczba miejsc w obiektach wczasowo-wypoczynkowych całorocznych

T2.7 liczba miejsc w obiektach wczasowo-wypoczynkowych w dniu 31.7 ogółem

T.2.8liczba miejsc w obiektach wczasowo-wypoczynkowych w dniu 31.7 całorocznych

3. Korzystający z noclegów

T3.1 liczba korzystających w obiektach turystycznych ogółem

T3.2 liczba korzystających w obiektach turystycznych całorocznych

T3.3 liczba korzystających w obiektach turystycznych cudzoziemców

T3.4 liczba korzystających w obiektach wczasowo-wypoczynkowych ogółem

T3.5 liczba korzystających w obiektach wczasowo-wypoczynkowych cudzoziemców

4. Udzielone noclegi

T4.1 liczba noclegów w obiektach turystycznych ogółem

T4.2 liczba noclegów w obiektach turystycznych całorocznych

T4.3 liczba noclegów w obiektach turystycznych (cudzoziemcy)

T4.4 nominalna liczba noclegów w obiektach turystycznych

T4.5 liczba noclegów w obiektach wczasowo-wypoczynkowych ogółem

T4.6 liczba noclegów w obiektach wczasowo-wypoczynkowych (cudzoziemcy)

T4.7 nominalna liczba noclegów w obiektach wczasowo-wypoczynkowych